

Computational Decision-Making Tools for Healthcare

Lead Guest Editor: Giovanni Improta

Guest Editors: Antonio Gloria, Massimo Martorelli, Paolo Gargiulo, and
Nuno Alves





Computational Decision-Making Tools for Healthcare

Journal of Healthcare Engineering

Computational Decision-Making Tools for Healthcare

Lead Guest Editor: Giovanni Improta

Guest Editors: Antonio Gloria, Massimo Martorelli,
Paolo Gargiulo, and Nuno Alves



Copyright © 2022 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Healthcare Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Associate Editors

Xiao-Jun Chen , China
Feng-Huei Lin , Taiwan
Maria Lindén, Sweden

Academic Editors

Cherif Adnen, Tunisia
Saverio Affatato , Italy
Óscar Belmonte Fernández, Spain
Sweta Bhattacharya , India
Prabadevi Boopathy , India
Weiwei Cai, USA
Gin-Shin Chen , Taiwan
Hongwei Chen, USA
Daniel H.K. Chow, Hong Kong
Gianluca Ciardelli , Italy
Olawande Daramola, South Africa
Elena De Momi, Italy
Costantino Del Gaudio , Italy
Ayush Dogra , India
Luobing Dong, China
Daniel Espino , United Kingdom
Sadiq Fareed , China
Mostafa Fatemi, USA
Jesus Favela , Mexico
Jesus Fontecha , Spain
Agostino Forestiero , Italy
Jean-Luc Gennisson, France
Badicu Georgian , Romania
Mehdi Gheisari , China
Luca Giancardo , USA
Antonio Gloria , Italy
Kheng Lim Goh , Singapore
Carlos Gómez , Spain
Philippe Gorce, France
Vincenzo Guarino , Italy
Muhammet Gul, Turkey
Valentina Hartwig , Italy
David Hewson , United Kingdom
Yan Chai Hum, Malaysia
Ernesto Iadanza , Italy
Cosimo Ieracitano, Italy

Giovanni Improta , Italy
Norio Iriguchi , Japan
Mihajlo Jakovljevic , Japan
Rutvij Jhaveri, India
Yizhang Jiang , China
Zhongwei Jiang , Japan
Rajesh Kaluri , India
Venkatachalam Kandasamy , Czech Republic
Pushpendu Kar , India
Rashed Karim , United Kingdom
Pasi A. Karjalainen , Finland
John S. Katsanis, Greece
Smith Khare , United Kingdom
Terry K.K. Koo , USA
Srinivas Koppu, India
Jui-Yang Lai , Taiwan
Kuruva Lakshmanna , India
Xiang Li, USA
Lun-De Liao, Singapore
Qiu-Hua Lin , China
Aiping Liu , China
Zufu Lu , Australia
Basem M. ElHalawany , Egypt
Praveen Kumar Reddy Maddikunta , India
Ilias Maglogiannis, Greece
Saverio Maietta , Italy
M.Sabarimalai Manikandan, India
Mehran Moazen , United Kingdom
Senthilkumar Mohan, India
Sanjay Mohapatra, India
Rafael Morales , Spain
Mehrbakhsh Nilashi , Malaysia
Sharnil Pandya, India
Jialin Peng , China
Vincenzo Positano , Italy
Saeed Mian Qaisar , Saudi Arabia
Alessandro Ramalli , Italy
Alessandro Reali , Italy
Vito Ricotta, Italy
Jose Joaquin Rieta , Spain
Emanuele Rizzuto , Italy

Dinesh Rokaya, Thailand
Sébastien Roth, France
Simo Saarakkala , Finland
Mangal Sain , Republic of Korea
Nadeem Sarwar, Pakistan
Emiliano Schena , Italy
Prof. Asadullah Shaikh, Saudi Arabia
Jiann-Shing Shieh , Taiwan
Tiago H. Silva , Portugal
Sharan Srinivas , USA
Kathiravan Srinivasan , India
Neelakandan Subramani, India
Le Sun, China
Fabrizio Taffoni , Italy
Jinshan Tang, USA
Ioannis G. Tollis, Greece
Ikram Ud Din, Pakistan
Sathishkumar V E , Republic of Korea
Cesare F. Valenti , Italy
Qiang Wang, China
Uche Wejinya, USA
Yuxiang Wu , China
Ying Yang , United Kingdom
Elisabetta Zanetti , Italy
Haihong Zhang, Singapore
Ping Zhou , USA

Contents

An Advance Computing Numerical Heuristic of Nonlinear SIR Dengue Fever System Using the Morlet Wavelet Kernel

Muhammad Umar, Zulqurnain Sabir, Muhammad Asif Zahoor Raja, K. S. Al-Basyouni, S. R. Mahmoud, and Yolanda Guerrero Sánchez 


Research Article (14 pages), Article ID 9981355, Volume 2022 (2022)

Application of DMAIC Cycle and Modeling as Tools for Health Technology Assessment in a University Hospital

Alfonso Maria Ponsiglione , Carlo Ricciardi , Arianna Scala , Antonella Fiorillo , Alfonso Sorrentino , Maria Triassi , Giovanni Dell'Aversana Orabona , and Giovanni Improta 





Research Article (11 pages), Article ID 8826048, Volume 2021 (2021)

A Machine Learning-Based Prediction Model for Preterm Birth in Rural India

Rakesh Raja , Indrajit Mukherjee, and Bikash Kanti Sarkar



Research Article (11 pages), Article ID 6665573, Volume 2021 (2021)

Big Data-Enabled Analysis of Factors Affecting Patient Waiting Time in the Nephrology Department of a Large Tertiary Hospital

Jialing Li , Guiju Zhu , Li Luo , and Wenwu Shen 



Research Article (10 pages), Article ID 5555029, Volume 2021 (2021)

A Hybrid Genetic Algorithm for Nurse Scheduling Problem considering the Fatigue Factor

Atefeh Amindoust , Milad Asadpour , and Samineh Shirmohammadi

Research Article (11 pages), Article ID 5563651, Volume 2021 (2021)

Involvement of Machine Learning Tools in Healthcare Decision Making

Senerath Mudalige Don Alexis Chinthaka Jayatilake  and Gamage Upeksha Ganegoda 

Review Article (20 pages), Article ID 6679512, Volume 2021 (2021)

Comparative Evaluation of the Treatment of COVID-19 with Multicriteria Decision-Making Techniques

Figen Sarigul Yildirim , Murat Sayan , Tamer Sanlidag, Berna Uzun, Dilber Uzun Ozsahin , and Ilker Ozsahin 


Research Article (11 pages), Article ID 8864522, Volume 2021 (2021)

Exploiting Multiple Optimizers with Transfer Learning Techniques for the Identification of COVID-19 Patients

Zeming Fan , Mudasir Jamil , Muhammad Tariq Sadiq , Xiwei Huang , and Xiaojun Yu 

Research Article (13 pages), Article ID 8889412, Volume 2020 (2020)

Surgical Design Optimization of Proximal Junctional Kyphosis

Li Peng, Guangming Zhang, Heng Zuo, Lan Lan , and Xiaobo Zhou

Research Article (8 pages), Article ID 8886599, Volume 2020 (2020)

Research Article

An Advance Computing Numerical Heuristic of Nonlinear SIR Dengue Fever System Using the Morlet Wavelet Kernel

Muhammad Umar,¹ Zulqurnain Sabir,¹ Muhammad Asif Zahoor Raja,² K. S. Al-Basyouni,³ S. R. Mahmoud,⁴ and Yolanda Guerrero Sánchez ⁵

¹Department of Mathematics and Statistics, Hazara University, Mansehra, Pakistan

²Future Technology Research Center, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan

³Mathematics Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

⁴GRC Department, Faculty of Applied Studies, King Abdulaziz University, Jeddah, Saudi Arabia

⁵Department of Dermatology, Stomatology, Radiology and Physical Medicine, University of Murcia, Murcia, Spain

Correspondence should be addressed to Yolanda Guerrero Sánchez; yolanda.guerreros@um.es

Received 11 March 2021; Accepted 5 January 2022; Published 31 January 2022

Academic Editor: Antonio Gloria

Copyright © 2022 Muhammad Umar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study is associated to solve the nonlinear SIR dengue fever system using a computational methodology by operating the neural networks based on the designed Morlet wavelet (MWNNs), global scheme as genetic algorithm (GA), and rapid local search scheme as interior-point algorithm (IPA), i.e., GA-IPA. The optimization of fitness function based on MWNNs is performed for solving the nonlinear SIR dengue fever system. This MWNNs-based fitness function is accessible using the differential system and initial conditions of the nonlinear SIR dengue fever system. The designed procedures based on the MWNN-GA-IPA are applied to solve the nonlinear SIR dengue fever system to check the exactness, precision, constancy, and efficiency. The achieved numerical form of the nonlinear SIR dengue fever system via MWNN-GA-IPA was compared with the Runge–Kutta numerical results that verify the significance of MWNN-GA-IPA. Moreover, statistical reflections through different measures for the nonlinear SIR dengue fever system endorse the precision and convergence of the computational MWNN-GA-IPA.

1. Introduction

Dengue fever disease (DFD) is one of the epidemics, infectious, and serious diseases that disturbed about 2.5 billion individuals all over the world. DFD occurred in some main countries of Southeast Asia due to the hot seasons. The infectious DFD grows fast, when the environment alters and becomes dangerous because of the shortage of information amongst the individuals [1]. DFD is an epidemiologic transmittable fever formed by dengue infection (DI) that is conveyed through mosquitoes to humans and apes [2]. Some symptoms scientifically produced by DFD are headache, joint pain, skin rash, and DI. The global World Health Organization (WHO) classified DI into a hemorrhagic

dengue fever before one decade [3]. DI is categorized into three organizational proteins, membrane protein (M), envelope protein (E), and capsid protein (C), while it has seven proteins that are nonstructural and their names are NS-I, NS-IIA, NS-III, NS-IIB, NS-IVA, NS-IVB, and NS-V [4]. According to envelope protein antigenicity, DI is ordered into four classes of microorganisms, which are infected as well as pathogenic [5].

DI has been reported in several zones of China and still any scientific report cannot prove it a confined endemic [6], whereas all other subgroups of DI have been imported; DI-I and DI-II are considered in China one of the main stereotypes endemic. In the Chinese province Guangdong, DV-I reported around 70% in 2014 and 80% in 2015. Presently,

an amino-based acid localizes transformation in NS-I of the Asian “Zika” inheritance, which has been described to grow the production of NS-I in the diseased host to make the virus convenient to Aedes type of mosquitoes, which is the focal wide spreading cause of virus “Zika” since 2015 [7].

Some epidemic infections such as COVID-19 are the existing transferred diseases, which have covered the whole world and the infection rate along with the number of deaths from COVID-19 increased steadily [8]. Another common disease is malaria, which is not directly transmitted from host to host. Protozoa is one of the transferred diseases, which spreads due to the anopheles of the female mosquito. According to the WHO report, almost 1/3 million individuals per year die from malaria. It distresses the children and pregnant women, mostly in the South African and American countries. Many chemical sprays have been widely implemented to control the mosquito population. Similarly, nonpolluting types of biological arrangements normally accomplished to emphasis the ecosystem of complicated kinds. The sterile insect apparatus (SIA) is a proficient nonpolluting scheme of insect control, which depends on the sterile male’s release. Consequently, adequately sterile

males discharging causes the elimination of the wild’s population. Over a half century ago, SIA has been recognized in the Curacao Island to adjust the screwworm hover [9, 10].

The aim of this study is to solve the nonlinear SIR dengue fever system using a computational methodology by operating the MWNNs, global search GA, and rapid local search IPA, i.e., MWNN-GA-IPA. The stochastic approaches have been investigated normally to solve a number of applications directed to the differential linear/nonlinear systems [11, 12]. However, no one has applied the MWNNs to solve the nonlinear SIR dengue fever system. Few recent reported submissions of stochastic solvers are eye surgery model, functional singular system, Thomas–Fermi singular equation, HIV-based infection system, biological form of the prey-predator system, periodic singular problems, singular three-point differential model, COVID-19 SITR system, multifractional singular models, system of heat transmission in human head, and mosquito spreading in heterogeneous conditions [13–15]. The intention of this study is to solve the nonlinear SIR dengue fever system using the MWNN-GA-IPA. The literature form of nonlinear SIR dengue fever system is written as [16]

$$\begin{cases} \frac{dX(\tau)}{d\tau} = \mu_h - \mu_h X(\tau) - \alpha X(\tau)Z(\tau), & X(0) = I_1, \\ \frac{dY(t)}{d\tau} = \alpha X(\tau)Z(\tau) - \beta Y(\tau), & Y(0) = I_2, \\ \frac{dZ(t)}{d\tau} = \gamma Y(\tau) - \gamma Y(\tau)Z(\tau) - \delta_1 Z(\tau), & Z(0) = I_3, \end{cases} \quad (1)$$

where the susceptible class, infected class, and recovered class are $X(\tau)$, $Y(\tau)$, and $Z(\tau)$, respectively. The terms γ , μ_h , δ_1 , α , and β used in system (1) are constant, whereas the initial conditions are I_1 , I_2 , and I_3 , respectively. Few major geographies of the MWNN-GA-IPA are concisely given as follows:

- (i) Design of Morlet wavelet is presented successfully as an activation function to solve the nonlinear SIR dengue fever system
- (ii) The reliable, consistent, and stable overlapped results obtained by the MWNN-GA-IPA and the true solutions validate the exactness of the proposed approach
- (iii) The authentication of the presentation is trained via different statistical valuations to get the solutions of the nonlinear SIR dengue fever system on multiple executions of the MWNN-GA-IPA

The rest of this paper is reported as follows: Section 2 indicates the proposed MWNN-GA-IPA along with the statistical measures. Section 3 shows the results simulations.

Section 4 describes the final remarks and future research reports.

2. Designed Procedure

The proposed structure of the MW-GA-IPA is used to solve the nonlinear SIR dengue fever system described in two phases as follows:

- (i) An objective function using the MW is considered to activate the neural networks
- (ii) Necessary clarifications are provided to enhance the merit function by applying the hybrid of GA-IPA

2.1. Designed Procedure Using MW Function. The mathematical design of the nonlinear SIR dengue fever system is described by using the achieved results of susceptible $\hat{X}(\tau)$, infected $\hat{Y}(\tau)$, and recovered $\hat{Z}(\tau)$ with the derivatives of these classes, written as [17]

$$\begin{aligned}
 [\widehat{X}(\tau), \widehat{Y}(\tau), \widehat{Z}(\tau)] &= \begin{bmatrix} \sum_{i=1}^m v_{X,i} H(w_{X,i}\tau + u_{X,i}), \sum_{i=1}^m v_{Y,i} H(w_{Y,i}\tau + u_{Y,i}), \\ \sum_{i=1}^m v_{Z,i} H(w_{Z,i}\tau + u_{Z,i}) \end{bmatrix}, \\
 [\widehat{X}^{(n)}(\tau), \widehat{Y}^{(n)}(\tau), \widehat{Z}^{(n)}(\tau)] &= \begin{bmatrix} \sum_{i=1}^m v_{X,i} H^{(n)}(w_{X,i}\tau + u_{X,i}), \sum_{i=1}^m v_{Y,i} H^{(n)}(w_{Y,i}\tau + u_{Y,i}), \\ \sum_{i=1}^m v_{Z,i} H^{(n)}(w_{Z,i}\tau + u_{Z,i}) \end{bmatrix},
 \end{aligned} \tag{2}$$

\mathbf{W} is the unknown weight vector given as $\mathbf{W} = [\mathbf{W}_X; \mathbf{W}_Y; \mathbf{W}_Z]$, for $\mathbf{W}_X = [v_X, \omega_X, \mathbf{u}_X]$, $\mathbf{W}_Y = [v_Y, \omega_Y, \mathbf{u}_Y]$, and $\mathbf{W}_Z = [v_Z, \omega_Z, \mathbf{u}_Z]$, where

$$\begin{aligned}
 \mathbf{v}_X &= [v_{X,1}; v_{X,2}; \dots; v_{X,m}], \quad \mathbf{v}_Y = [v_{Y,1}; v_{Y,2}; \dots; v_{Y,m}], \quad \mathbf{v}_Z = [v_{Z,1}; v_{Z,2}; \dots; v_{Z,m}], \\
 \mathbf{w}_X &= [w_{X,1}; w_{X,2}; \dots; w_{X,m}], \quad \mathbf{w}_Y = [w_{Y,1}; w_{Y,2}; \dots; w_{Y,m}], \quad \mathbf{w}_Z = [w_{Z,1}; w_{Z,2}; \dots; w_{Z,m}], \\
 \mathbf{u}_X &= [u_{X,1}; u_{X,2}; \dots; u_{X,m}], \quad \mathbf{u}_Y = [u_{Y,1}; u_{Y,2}; \dots; u_{Y,m}], \quad \mathbf{u}_Z = [u_{Z,1}; u_{Z,2}; \dots; u_{Z,m}].
 \end{aligned} \tag{3}$$

The MWNN has not been implemented before to solve the nonlinear SIR dengue fever system. The mathematical form of MW function is written as [18]

$$H(\tau) = \cos(1.75)e^{-0.5\tau^2}. \tag{4}$$

The simplified form of system (2) using the above MW function is given as

$$\begin{aligned}
 [\widehat{X}(\tau), \widehat{Y}(\tau), \widehat{Z}(\tau)] &= \begin{bmatrix} \sum_{i=1}^m v_{X,i} \cos(1.75(w_{X,i}\tau + u_{X,i}))e^{-0.5(w_{X,i}\tau + u_{X,i})^2}, \\ \sum_{i=1}^m v_{Y,i} \cos(1.75(w_{Y,i}\tau + u_{Y,i}))e^{-0.5(w_{Y,i}\tau + u_{Y,i})^2}, \\ \sum_{i=1}^m v_{Z,i} \cos(1.75(w_{Z,i}\tau + u_{Z,i}))e^{-0.5(w_{Z,i}\tau + u_{Z,i})^2}, \end{bmatrix}, \\
 [\widehat{X}^{(n)}(\tau), \widehat{Y}^{(n)}(\tau), \widehat{Z}^{(n)}(\tau)] &= \frac{d}{d\tau} \begin{bmatrix} \sum_{i=1}^m v_{X,i} \cos(1.75(w_{X,i}\tau + u_{X,i}))e^{-0.5(w_{X,i}\tau + u_{X,i})^2}, \\ \sum_{i=1}^m v_{Y,i} \cos(1.75(w_{Y,i}\tau + u_{Y,i}))e^{-0.5(w_{Y,i}\tau + u_{Y,i})^2}, \\ \sum_{i=1}^m v_{Z,i} \cos(1.75(w_{Z,i}\tau + u_{Z,i}))e^{-0.5(w_{Z,i}\tau + u_{Z,i})^2}. \end{bmatrix}.
 \end{aligned} \tag{5}$$

An error function based on the merit function is given as

$$e = \sum_{j=1}^4 e_j, \quad (6)$$

$$e_1 = \frac{1}{N} \sum_{i=1}^N \left[\widehat{X}'_i - \mu_h + \mu_h \widehat{X}_i + \alpha \widehat{X}_i \widehat{Z}_i \right]^2, \quad (7)$$

$$e_2 = \frac{1}{N} \sum_{i=1}^N \left[\widehat{Y}'_i - \alpha \widehat{X}_i \widehat{Z}_i + \beta \widehat{Y}_i \right]^2, \quad (8)$$

$$e_3 = \frac{1}{N} \sum_{i=1}^N \left[\widehat{Z}'_i - \gamma \widehat{Y}_i + \gamma \widehat{Z}_i \widehat{Y}_i + \delta_1 \widehat{Z}_i \right]^2, \quad (9)$$

$$e_4 = \frac{1}{3} \left[(\widehat{X}_0 - I_1)^2 + (\widehat{Y}_0 - I_2)^2 + (\widehat{Z}_0 - I_3)^2 \right], \quad (10)$$

where $\widehat{X}_i = X(\tau_i)$, $\widehat{Y}_i = Y(\tau_i)$, $\widehat{Z}_i = Z(\tau_i)$, $Nh = 1$, and $\tau_i = ih$. \widehat{X}_i , \widehat{Y}_i and \widehat{Z}_i show the proposed results of the susceptible class, infected class, and recovered class. Likewise, e_1, e_2 , and e_3 denote the error function related to system (1), whereas e_4 denotes the error function on the basis of initial conditions.

2.2. Optimization: MWNN-GA-IPA. The optimization performance is presented for solving the nonlinear SIR dengue fever system using the MWNN-GA-IPA. The structure of the present approach to solve the nonlinear SIR dengue fever system is provided in Figure 1.

GA is a global optimization procedure, which is executed to solve the nonlinear SIR dengue fever system by implementing the usual selection procedures. GA is pragmatic frequently to regulate the accurate population to solve several

complicated or stiff systems. To attain the best model outcomes, GAs operate through the operators based on selection, reproduction, crossover, and mutation. Few existing GA's applications are the hospitalization expenditure system [19], feature assortment in cancer microarray [20], organization of irregular magnetic character brain tumor imageries [21], vehicle routing system [22], prediction-based traffic flow system [23], radiation shielding optimizations in the bismuth-borate spectacles [24], prediction of air blast [25], composition optimization of cloud service [26], task arrangement models in phased range radar [27], arrangement system of microarray cancer [28], system dynamics of monorail vehicle [17], and prediction system of liver disease [29].

IPA is known as an optimized local search approach, which is performed broadly in both types of models (constrained/unconstrained). IPA is used in the optimization of various complicated and nonstiff natured systems. Recently, IPA is executed for image restoration [30], multistage nonlinear nonconvex models [31], viscoplastic fluid flows [32], nonsmooth contact dynamics [33], power systems [34], and dynamic flux balance analysis models [35]. The hybridization process of GA-IPA is applied to remove the laziness of GA, i.e., global approach. The pseudocode based on the designed approach MWNN-GA-IPA is provided in Table 1.

2.3. Performance Measures. The mathematical measures using the statistical operators for variance accounted for (VAF), semi-interquartile (S.I) range, Theil's inequality coefficient (T.I.C), and mean absolute deviation (M.A.D) along with the Global VAF (G-VAF), Global M.A.D (G-M.A.D), and Global T.I.C to solve the nonlinear SIR dengue fever system which is given as

$$\left\{ \begin{array}{l} [\text{V.A.F}_X, \text{V.A.F}_Y, \text{V.A.F}_Z] = \left[\begin{array}{l} \left(1 - \frac{\text{var}(X_r - \widehat{X}_r)}{\text{var}(X_r)} \right) * 100, \left(1 - \frac{\text{var}(Y_r - \widehat{Y}_r)}{\text{var}(Y_r)} \right) * 100, \\ \left(1 - \frac{\text{var}(Z_r - \widehat{Z}_r)}{\text{var}(Z_r)} \right) * 100, \end{array} \right], \\ [\text{E} - \text{V.A.F}_X, \text{E} - \text{V.A.F}_Y, \text{E} - \text{V.A.F}_Z] = [100 - \text{V.A.F}_X, 100 - \text{V.A.F}_Y, 100 - \text{V.A.F}_Z]. \end{array} \right. \quad (11)$$

$$\left\{ \begin{array}{l} \text{S.I Range} = -0.5 \times (Q_1 - Q_3), \\ Q_1 = 1^{\text{st}} \text{ quartile} \ \& \ Q_3 = 3^{\text{rd}} \text{ quartile}, \end{array} \right. \quad (12)$$

$$[\text{T.I.C}_X, \text{T.I.C}_Y, \text{T.I.C}_Z] = \left[\begin{array}{l} \frac{\sqrt{(1/n) \sum_{r=1}^n (X_r - \widehat{X}_r)^2}}{\left(\sqrt{(1/n) \sum_{r=1}^n X_r^2} + \sqrt{(1/n) \sum_{r=1}^n \widehat{X}_r^2} \right)}, \frac{\sqrt{(1/n) \sum_{r=1}^n (Y_r - \widehat{Y}_r)^2}}{\left(\sqrt{(1/n) \sum_{r=1}^n Y_r^2} + \sqrt{(1/n) \sum_{r=1}^n \widehat{Y}_r^2} \right)}, \\ \frac{\sqrt{(1/n) \sum_{r=1}^n (Z_r - \widehat{Z}_r)^2}}{\left(\sqrt{(1/n) \sum_{r=1}^n Z_r^2} + \sqrt{(1/n) \sum_{r=1}^n \widehat{Z}_r^2} \right)}, \end{array} \right] \quad (13)$$

$$[\text{M.A.D}_X, \text{M.A.D}_Y, \text{M.A.D}_Z] = \left[\sum_{r=1}^n |X_r - \widehat{X}_r|, \sum_{r=1}^n |Y_r - \widehat{Y}_r|, \sum_{r=1}^n |Z_r - \widehat{Z}_r| \right],$$

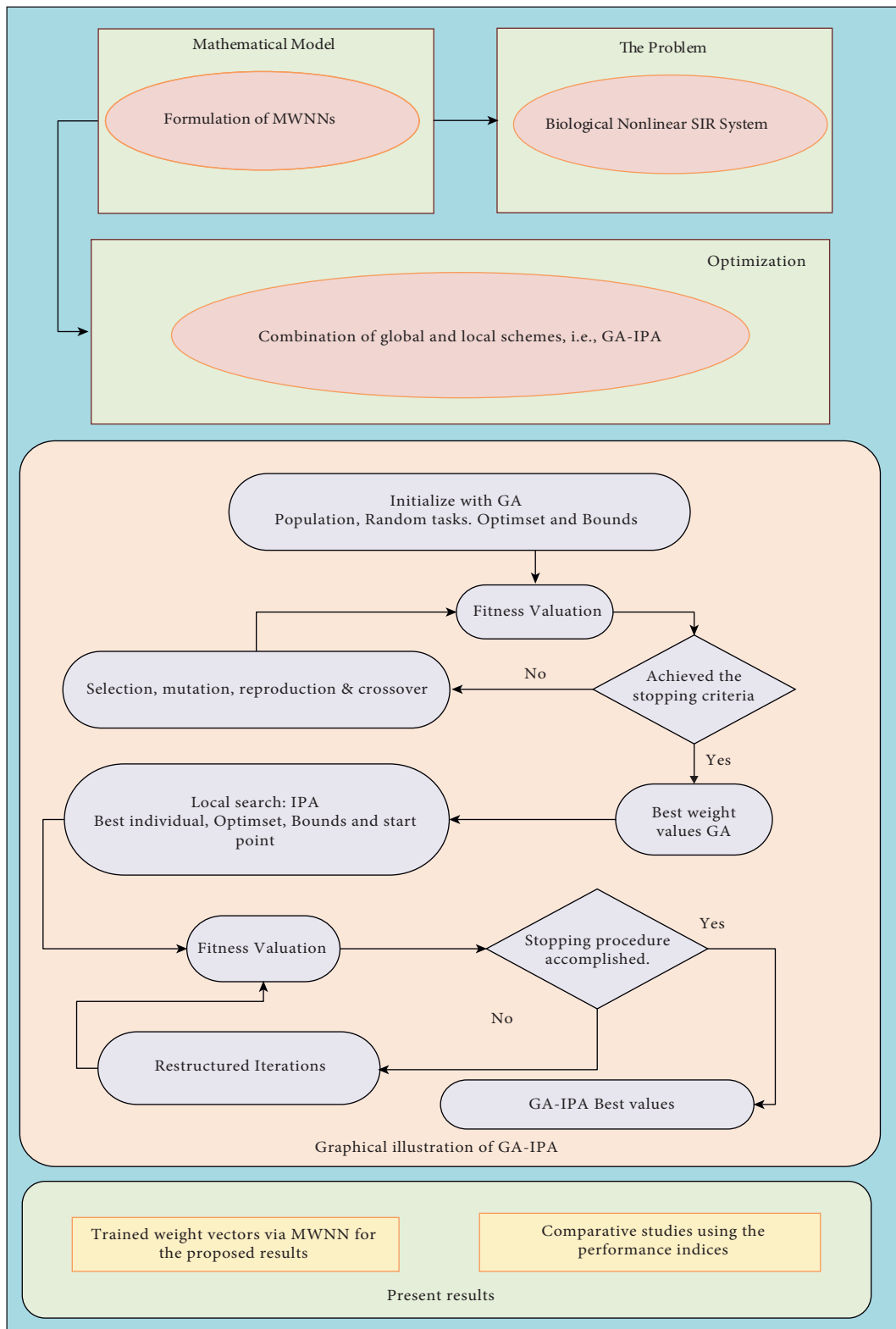


FIGURE 1: Structure of the present approach to solve the nonlinear SIR dengue fever system.

TABLE 1: Optimization performance taking the MWNN-GA-IPA for the nonlinear SIR dengue fever system.

Start of GA

Inputs: the chromosomes are characterized with the same system element as

$$W = [v, w, u]$$

Population: the chromosomes set is written as

$$W_X = [v_X, \omega_X, u_X], W_Y = [v_Y, \omega_Y, u_Y] \text{ and}$$

Output: global values of the weight are represented as $W_{GA-Best}$

Initialization: for the selection of chromosomes, select the weight vector values.

Fit evaluation: modify the values of fitness “ e ” in population “ P ” for each vector with the use of systems 4–8

(i) **Stopping criteria:** terminate when [$e = 10^{-21}$], [Generations = 55], [StallLimit = 140], [PopSize = 285], and [TolFun = TolCon = 10^{-21}]

Move to **storage**

Ranking: rank individual weight vector in population using the values of the fitness

Storage: save $W_{GA-Best}$, iterations, time, e , and count of function for the presence of GA

End of GA**IPA starts**

Inputs: start point: $W_{GA-Best}$

Output: W_{GA-IPA} shows the best weight values of GA-IPA

Initialize: $W_{GA-Best}$, iterations, assignments, and other values

Terminate: stop, when [$e = 10^{-20}$], [Iterations = 750], [MaxFunEvals = 267000], [TolCon = TolX = 10^{-22}], and [TolFun = 10^{-22}] achieved.

Evaluation of fitness: compute W and e using equations (8)–(12)

Amendments: adjust “fmincon” for IPA, compute e of better-quality of ‘ W ’ using systems 4–8

Accumulate: transmute W_{GA-IPA} , e , function counts, iterations, and time for the existing IPA runs

IPA process ends

where the approximate solutions are \hat{X} , \hat{Y} , and \hat{Z} , respectively.

3. Simulations of the Results

The current work is associated to solve the nonlinear SIR dengue fever system shown in system (1). The relative

presentation of the obtained results using the Runge–Kutta solutions is tested to form the correctness of MWNN-GA-IPA. Additionally, statistical operators indicate the precision and accuracy of MWNN-GA-IPA. The simplified measures of the nonlinear SIR dengue fever system using the suitable values are given as

$$\begin{cases} \hat{X}'(\tau) = 0.000046 - (0.000046 + 0.375Z(\tau))X(\tau), & X(0) = 0.9999, \\ \hat{Y}'(\tau) = 0.375X(\tau)Z(\tau) - 0.0323Y(\tau) & Y(0) = 0.0006, \\ \hat{Z}'(\tau) = 0.328833 - (0.328833Y(\tau) + 0.0001)Z(\tau), & Z(0) = 0.0560. \end{cases} \quad (14)$$

The fitness function for system (14) is written as

$$\begin{aligned} e = \frac{1}{N} \sum_{i=1}^N & \left(\left[\hat{X}'_r - (0.000046 + 0.375\hat{Z}_r)\hat{X}_r \right]^2 + \left[\hat{Y}'_r - 0.375\hat{X}_r\hat{Z}_r + 0.0323\hat{Y}_r \right]^2 \right) \\ & + \left[\hat{Z}'_r - (0.328833\hat{Y}_r + 0.0001)\hat{Z}_r \right]^2 \\ & + \frac{1}{3} \left[\left(\hat{X}_0 - \frac{9999}{10000} \right)^2 + \left(\hat{Y}_0 - \frac{6}{10000} \right)^2 + \left(\hat{Z}_0 - \frac{56}{1000} \right)^2 \right]. \end{aligned} \quad (15)$$

The nonlinear SIR dengue fever system given in system (1) is optimized using the MWNN-GA-IPA for 100 trials to attain ANN model parameters for 10 neurons. Figure 1 is

drawn using the best outputs of the weight vector, i.e., W for the MWNN-GA-IPA. These best weights of the output are applied to solve the estimated outcomes of the nonlinear SIR

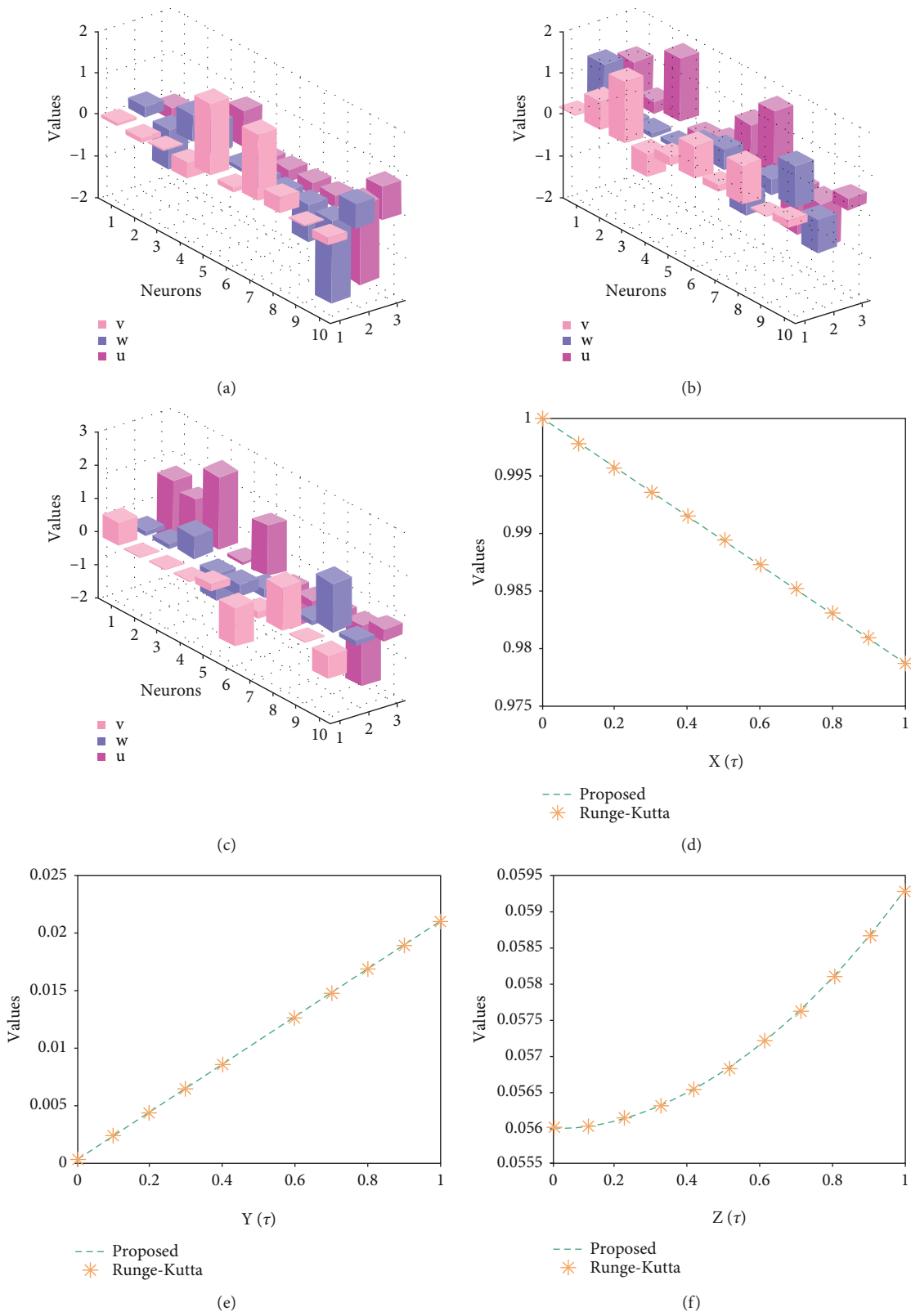


FIGURE 2: Best weight vector set and result comparison for each class of nonlinear SIR dengue fever system. (a) Best weights of $X(\tau)$ for 10 neurons. (b) Best weights of $Y(\tau)$ for 10 neurons. (c) Best weights of $Z(\tau)$ for 10 neurons. (d) Comparison for $X(\tau)$ class. (e) Comparison for $Y(\tau)$ class. (f) Comparison for $Z(T)$ class.

dengue fever system. The mathematical illustrations of these estimated results from MWNN-GA-IPA are given as

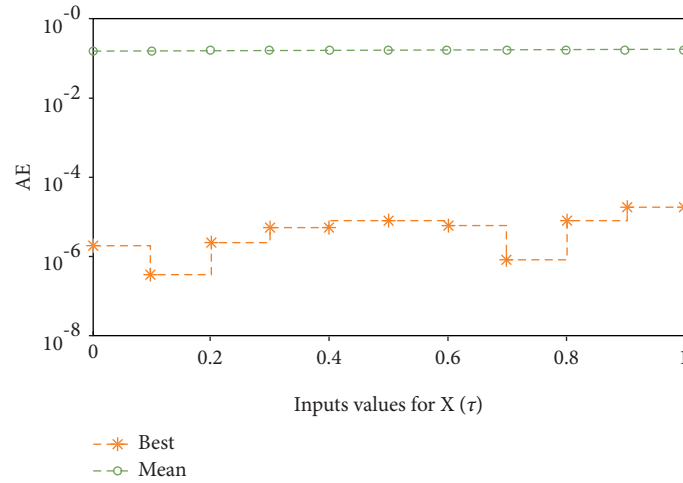
$$\begin{aligned} \widehat{X}(\tau) = & 0.067 \cos(1.75(0.2561\tau - 0.2972))e^{-0.5(0.2561\tau - 0.2972)^2} \\ & - 0.1262 \cos(1.75(-0.957\tau - 0.6639))e^{-0.5(-0.957\tau - 0.6639)^2} \\ & + 0.0208 \cos(1.75(0.6513\tau - 0.0050))e^{-0.5(0.6513\tau - 0.0050)^2} \\ & - 0.3607 \cos(1.75(0.7574\tau - 0.8275))e^{-0.5(0.7574\tau - 0.8275)^2} \\ & + 1.7204 \cos(1.75(-0.058\tau - 0.3235))e^{-0.5(-0.058\tau - 0.3235)^2} \\ & - 0.1039 \cos(1.75(-0.271\tau - 0.5232))e^{-0.5(-0.271\tau - 0.5232)^2} \\ & + 1.5608 \cos(1.75(-0.3176\tau - 1.242))e^{-0.5(-0.3176\tau - 1.242)^2} \\ & + 0.3663 \cos(1.75(-0.8974\tau - 0.693))e^{-0.5(-0.8974\tau - 0.693)^2} \\ & - 0.0041 \cos(1.75(-2.0010\tau - 1.863))e^{-0.5(-2.0010\tau - 1.863)^2} \\ & - 0.1883 \cos(1.75(0.61200\tau + 0.813))e^{-0.5(0.61200\tau + 0.813)^2}, \end{aligned} \quad (16)$$

$$\begin{aligned} \widehat{Y}(\tau) = & 0.0097 \cos(1.75(1.0108\tau + 0.8687))e^{-0.5(1.0108\tau + 0.8687)^2} \\ & + 0.6808 \cos(1.75(-0.475\tau + 0.1533))e^{-0.5(-0.475\tau + 0.1533)^2} \\ & + 1.3976 \cos(1.75(-0.107\tau + 1.5804))e^{-0.5(-0.107\tau + 1.5804)^2} \\ & - 0.5523 \cos(1.75(-0.066\tau - 0.3244))e^{-0.5(-0.066\tau - 0.3244)^2} \\ & + 0.2293 \cos(1.75(0.3343\tau + 0.1346))e^{-0.5(0.3343\tau + 0.1346)^2} \\ & + 0.8040 \cos(1.75(0.5024\tau + 0.8675))e^{-0.5(0.5024\tau + 0.8675)^2} \\ & + 0.1466 \cos(1.75(-0.761\tau + 1.5040))e^{-0.5(-0.761\tau + 1.5040)^2} \\ & + 0.9269 \cos(1.75(0.3945\tau - 1.1415))e^{-0.5(0.3945\tau - 1.1415)^2} \\ & + 0.0374 \cos(1.75(0.9908\tau - 1.2051))e^{-0.5(0.9908\tau - 1.2051)^2} \\ & + 0.1970 \cos(1.75(-0.764\tau + 0.2823))e^{-0.5(-0.764\tau + 0.2823)^2}, \end{aligned} \quad (17)$$

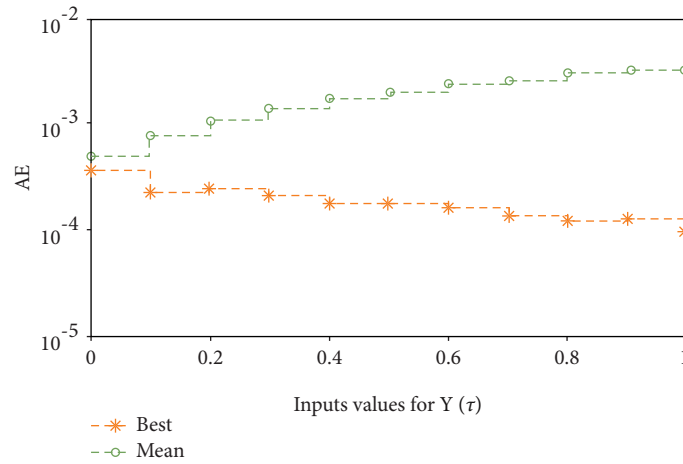
$$\begin{aligned} \widehat{Z}(\tau) = & 0.607 \cos(1.75(0.1112\tau + 1.3610))e^{-0.5(0.1112\tau + 1.3610)^2} \\ & - 0.0384 \cos(1.75(0.1343\tau + 1.1398))e^{-0.5(0.1343\tau + 1.1398)^2} \\ & + 0.0150 \cos(1.75(0.7727\tau + 2.2949))e^{-0.5(0.7727\tau + 2.2949)^2} \\ & - 0.0024 \cos(1.75(-0.829\tau + 0.0417))e^{-0.5(-0.829\tau + 0.0417)^2} \\ & + 0.3144 \cos(1.75(-0.365\tau + 1.4552))e^{-0.5(-0.365\tau + 1.4552)^2} \\ & - 1.1990 \cos(1.75(0.2513\tau - 1.1402))e^{-0.5(0.2513\tau - 1.1402)^2} \\ & + 0.1784 \cos(1.75(0.0640\tau - 0.1835))e^{-0.5(0.0640\tau - 0.1835)^2} \\ & + 1.3200 \cos(1.75(0.0134\tau - 0.5702))e^{-0.5(0.0134\tau - 0.5702)^2} \\ & - 0.0002 \cos(1.75(1.5840\tau - 1.8226))e^{-0.5(1.5840\tau - 1.8226)^2} \\ & - 0.8039 \cos(1.75(0.2124\tau + 0.2349))e^{-0.5(0.2124\tau + 0.2349)^2}. \end{aligned} \quad (18)$$

Systems (16)–(18) are implemented to solve the non-linear SIR dengue fever system given in system (1) using the MWNN-GA-IPA and the acquired results are plotted in

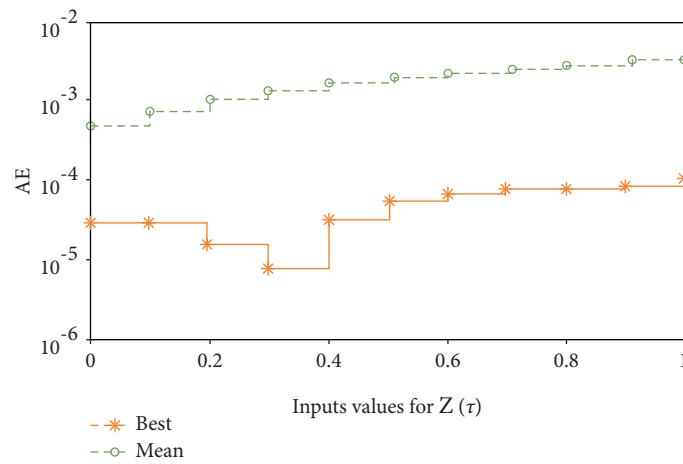
Figures 2–4. Figure 2 shows the set of best weights and comparison of the best obtained results with the Runge–Kutta numerical results. It is seen that the proposed and



(a)



(b)



(c)

FIGURE 3: AE values for each class of the nonlinear SIR dengue fever system. (a) AE for $X(\tau)$ class. (b) AE for $Y(\tau)$ class. (c) AE for $Z(\tau)$ class.

reference results overlapped each other for $\hat{X}(\tau)$, $\hat{Y}(\tau)$, and $\hat{Z}(\tau)$ classes to solve the nonlinear SIR dengue fever system. The plots of the AE for $\hat{X}(\tau)$, $\hat{Y}(\tau)$, and $\hat{Z}(\tau)$ classes to solve the nonlinear SIR dengue fever system are reported in

Figure 3. For the $\hat{X}(\tau)$ class, $\hat{Y}(\tau)$ class, and $\hat{Z}(\tau)$ class, the AE best values lie about 10^{-6} - 10^{-8} , 10^{-3} - 10^{-5} , and 10^{-4} - 10^{-6} , and the AE mean values lie around 10^{-1} - 10^{-2} , 10^{-2} - 10^{-3} , and 10^{-3} - 10^{-4} , respectively. The performance

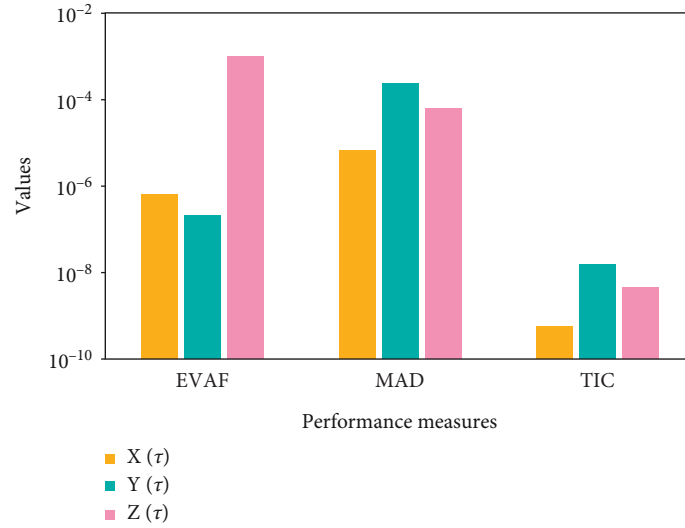
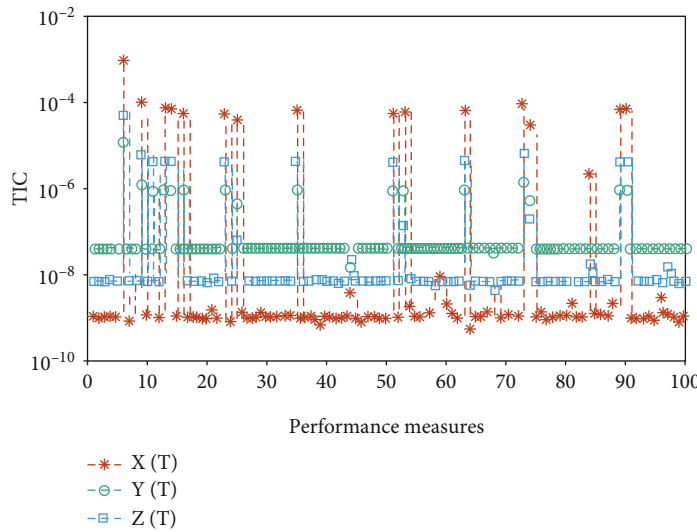
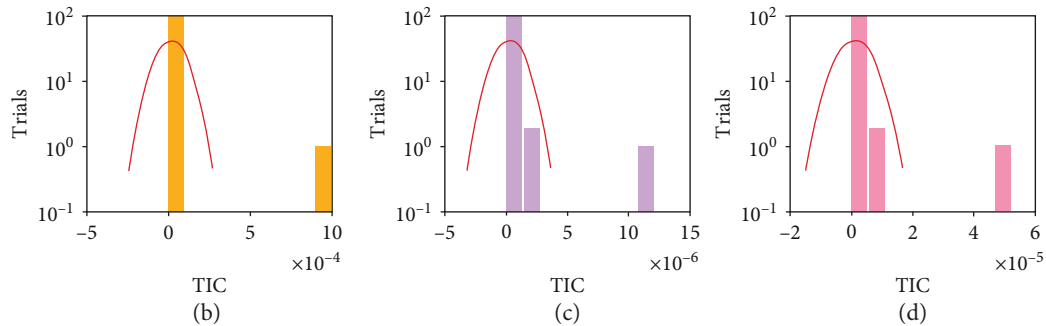


FIGURE 4: Performance of the E-VAF, M.A.D, and T.I.C operators for solving each class of the nonlinear SIR dengue fever system. (a) Performance for each class of the nonlinear SIR system.



(a)



(b)

(c)

(d)

FIGURE 5: Convergence of T.I.C plots along with the histogram using MWNN-GA-IPA to solve each class of the nonlinear SIR dengue fever system. (a) T.I.C for each class of the nonlinear SIR system. (b) Histogram for $X(\tau)$ class. (c) Histogram for $Y(\tau)$ class. (d) Histogram for $Z(\tau)$ class.

plots of the E-VAF and T.I.C indices for each class of the nonlinear SIR dengue fever system are plotted in Figure 4. For $\hat{X}(\tau)$ category, the best E-VAF, M.A.D, and T.I.C values

lie around 10⁻⁶-10⁻⁸, 10⁻⁵-10⁻⁶, and 10⁻⁹-10⁻¹⁰. For $\hat{Y}(\tau)$ category, the best E-VAF, M.A.D, and T.I.C values lie around 10⁻⁵-10⁻⁶, 10⁻³-10⁻⁴, and 10⁻⁸-10⁻⁹. Similarly, for

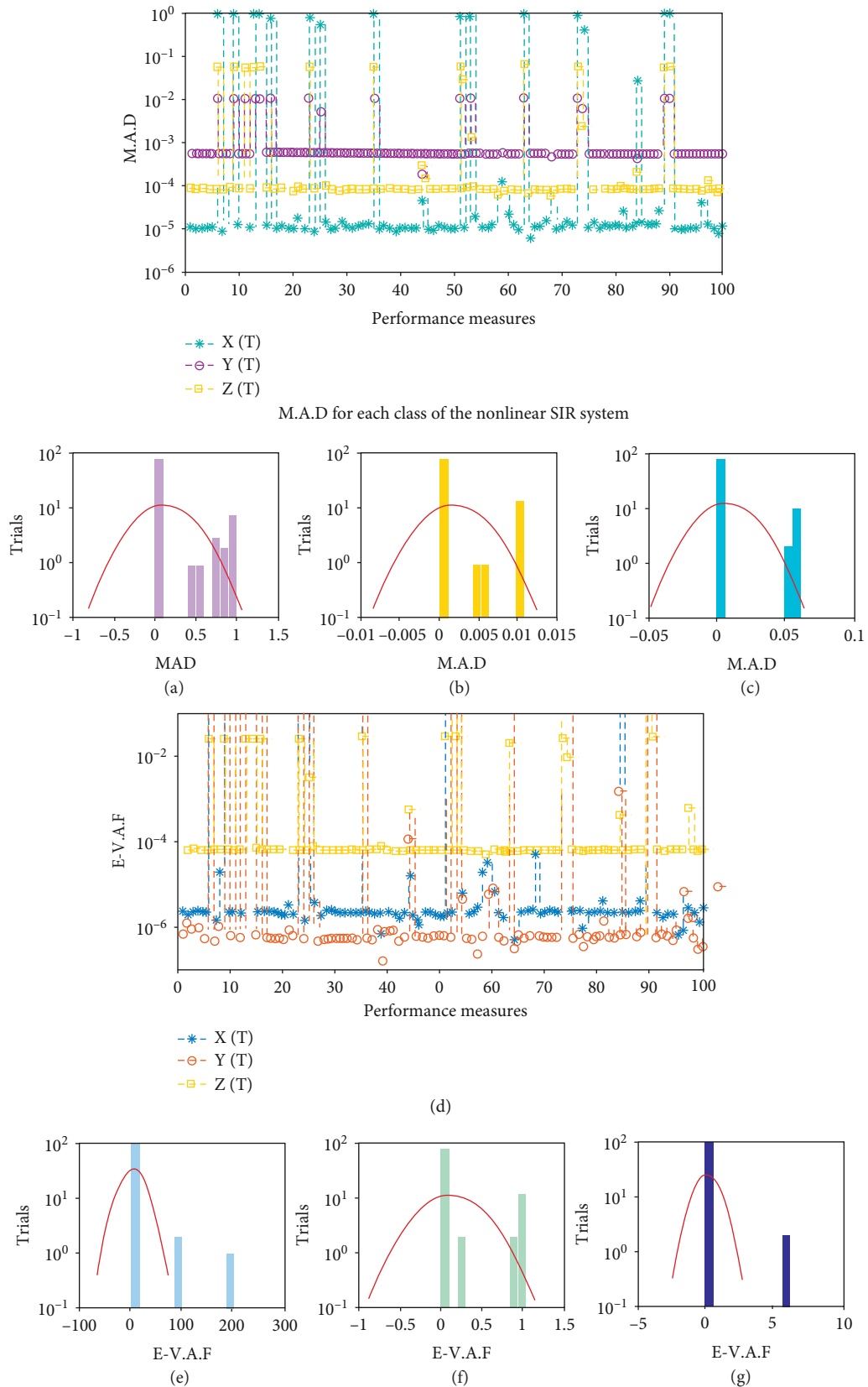


FIGURE 6: Convergence of M.A.D and E-V.A.F plots along with the histogram using MWNN-GA-IPA to solve each class of the nonlinear SIR dengue fever system. (a) Histogram for $X(\tau)$ class. (b) Histogram for $Y(\tau)$ class. (c) Histogram for $Z(\tau)$ class. (d) E-V.A.F for each class of the nonlinear SIR system. (e) Histogram for $X(\tau)$ class. (f) Histogram for $Y(\tau)$ class. (g) Histogram for $Z(\tau)$ class.

TABLE 2: Statistical presentations of the nonlinear SIR dengue fever system for the category $X(\tau)$.

τ	$X(\tau)$				
	Min	Max	Median	S.I range	S.T.D
0	4.0423220E-12	9.9990000E-01	2.2132198E-07	3.3825540E-06	2.9955460E-01
0.1	1.8562196E-08	9.9800585E-01	2.2575664E-06	9.2252690E-06	3.1917267E-01
0.2	1.1989577E-08	9.9570595E-02	2.8147596E-06	7.3089931E-06	3.1842462E-02
0.3	3.2594992E-07	9.9360942E-01	3.6204945E-06	5.6263588E-06	3.1769837E-01
0.4	9.7767938E-08	9.9151010E-01	5.4797252E-06	5.1357097E-06	3.1697503E-01
0.5	3.2511304E-06	9.8940563E-02	8.1700536E-06	7.3575618E-06	3.1625439E-02
0.6	2.8132523E-06	9.8729376E-01	1.1797241E-05	1.1757718E-05	3.1554735E-01
0.7	6.9767773E-07	9.8517226E-01	1.5987828E-05	9.1401060E-06	3.1478449E-01
0.8	6.3270889E-06	9.8303897E-02	2.0912004E-05	1.0215585E-05	3.1411641E-02
0.9	4.5794286E-06	9.8089357E-01	2.5581289E-05	9.6963771E-06	3.1339004E-01
1	3.0336439E-06	9.7981727E-01	3.1364852E-05	1.1245408E-05	3.1271044E-02

TABLE 3: Statistical presentations of the nonlinear SIR dengue fever system for the category $Y(\tau)$.

τ	$Y(\tau)$				
	Min	Max	Median	S.I range	S.T.D
0	4.7826156E-05	1.2441636E-03	5.3999940E-04	2.1685104E-07	1.9190429E-04
0.1	2.2617648E-04	3.1953335E-03	5.3868377E-04	2.2227911E-06	6.0207993E-04
0.2	2.4551957E-04	5.1214677E-03	5.3800062E-04	2.9618951E-06	1.3159469E-03
0.3	2.1066905E-04	7.0066945E-03	5.3756084E-04	2.1655938E-06	2.0394513E-03
0.4	1.7704587E-04	8.9185145E-03	5.3809794E-04	1.9832062E-06	2.7655709E-03
0.5	1.7527548E-04	1.0862314E-02	5.3919348E-04	1.6099165E-06	3.4944298E-03
0.6	1.5351687E-04	1.2834137E-02	5.4066226E-04	1.7106495E-06	4.2264704E-03
0.7	4.2598412E-05	1.4831228E-02	5.4286714E-04	2.3156731E-06	4.9614134E-03
0.8	1.2432002E-04	1.6851840E-02	5.4585539E-04	3.3573941E-06	5.6986905E-03
0.9	1.2880654E-04	1.8894845E-02	5.4928063E-04	3.3031846E-06	6.4404297E-03
1	9.7394166E-05	2.0992829E-02	5.5344362E-04	2.9146189E-06	7.1876800E-03

TABLE 4: Statistical presentations of the nonlinear SIR dengue fever system for the category $Z(\tau)$.

τ	$Z(\tau)$				
	Min	Max	Median	S.I range	S.T.D
0	2.7359906E-11	6.0293074E-02	2.1020438E-07	1.4892650E-06	1.6947015E-02
0.1	1.2779827E-06	6.0606789E-02	1.7227698E-05	2.9467734E-06	1.8379946E-02
0.2	9.8289161E-07	6.0932383E-02	3.4063897E-05	4.1260050E-06	1.8423547E-02
0.3	7.8962744E-06	6.1220946E-02	5.0622682E-05	2.7004041E-06	1.8477633E-02
0.4	2.8442743E-05	6.1487380E-02	6.6956094E-05	1.7144321E-06	1.8548676E-02
0.5	4.3296911E-05	6.1745495E-02	8.3558913E-05	2.8627209E-06	1.8637956E-02
0.6	6.4428148E-05	6.2007841E-02	1.0025700E-04	3.3723574E-06	1.8746689E-02
0.7	7.8570024E-05	6.2285742E-02	1.1721379E-04	3.3496552E-06	1.8875847E-02
0.8	8.0198804E-05	6.2589332E-02	1.3406477E-04	3.9024126E-06	1.9025869E-02
0.9	4.5991286E-05	6.2927739E-02	1.5084468E-04	2.3165808E-06	1.9198331E-02
1	4.4797926E-05	6.3311235E-02	1.6779808E-04	4.5753017E-06	1.9389410E-02

$\widehat{Z}(\tau)$ class, the best E-VAF, M.A.D, and T.I.C values lie around 10^{-3} - 10^{-4} , 10^{-4} - 10^{-5} , and 10^{-9} - 10^{-10} .

The graphical representations of the statistical trials along with the values of histograms are shown in Figures 5 and 6 for each class of nonlinear SIR dengue fever system. The convergence based on the E-VAF, M.A.D, and T.I.C operators is accomplished for independent trials to the nonlinear SIR dengue fever system. The achieved results from MWNN-GA-IPA are calculated satisfactory based on the T.I.C, M.A.D, and E-VAF operators.

For the accurateness and precision measures, statistical studies are provided in Tables 2-4 to solve each class of the nonlinear SIR dengue fever using the operatives minimum (Min), S.I range, maximum (Max), standard deviation (S.T.D), and median. The Min and Max standards show the best results and poorest results in the 100 executions. For $X(\tau)$ category, the Min, Max, median, S.I range, and S.T.D values lie around 10^{-7} - 10^{-12} , 10^{-1} - 10^{-2} , 10^{-5} - 10^{-7} , 10^{-5} - 10^{-6} , and 10^{-1} - 10^{-2} , respectively. For the category $Y(\tau)$, the Min, Max, median, S.I range, and S.T.D values lie around

TABLE 5: Global presentations for each category of the nonlinear SIR dengue fever system.

Category	[G-M.A.D]		[G-T.I.C]		[G-E.VAF]	
	Mean	S.I range	Mean	S.I range	Mean	S.I range
$X(\tau)$	$1.14673E-05$	$7.46457E-06$	$1.09132E-09$	$5.61892E-11$	$2.30930E-06$	$1.16918E-06$
$Y(\tau)$	$5.42141E-04$	$1.85278E-07$	$3.86220E-08$	$1.58454E-10$	$6.26006E-07$	$4.76894E-07$
$Z(\tau)$	$8.40963E-05$	$2.85409E-06$	$7.05895E-09$	$2.03526E-10$	$6.46250E-05$	$1.58215E-06$

10^{-4} - 10^{-5} , 10^{-2} - 10^{-3} , 10^{-4} - 10^{-5} , 10^{-6} - 10^{-7} , and 10^{-3} - 10^{-4} , respectively. Likewise, the Min, Max, median, S.I range, and S.T.D values for the category $Z(\tau)$ lie around 10^{-5} - 10^{-11} , 10^{-2} - 10^{-3} , 10^{-4} - 10^{-7} , 10^{-6} - 10^{-7} , and 10^{-2} - 10^{-3} , respectively. These calculated presentations found the worth and value of the proposed MWNN-GA-IPA to solve the nonlinear SIR dengue fever system. One can establish through the achieved results that the MWNN-GA-IPA is stable and precise.

The global performance of the operators [G-M.A.D], [G-T.I.C], and [G-E.VAF] for 100 trials of MWNN-GA-IPA is plotted in Table 5 to solve each category of the nonlinear SIR dengue fever system. The global-based mean [G-M.A.D], [G-T.I.C], and [G-E.VAF] values are found to be 10^{-4} - 10^{-5} , 10^{-8} - 10^{-9} , and 10^{-5} - 10^{-7} , whereas the global values of the S.I Range lie in the interval 10^{-6} - 10^{-7} , 10^{-10} - 10^{-11} , and 10^{-6} - 10^{-7} for each category of the nonlinear SIR dengue fever system. The close optimal outcomes acquired by the global measures approve the accurateness, correctness, and precision of MWNN-GA-IPA.

4. Conclusions

The current investigations are linked to design a neural network based on Morlet wavelet (MWNN) function for solving the nonlinear SIR dengue fever system based on dengue infection using the optimization procedures of global and local search approaches, i.e., GA-IPA. The nonlinear SIR dengue fever system is capable to evaluate through GA-IPA using the layer arrangement of Morlet wavelet neural networks taking 10 neurons. The overlapped results through MWNN-GA-IPA and the reference results show the good accuracy level to solve the nonlinear SIR dengue fever system based on dengue infection. The performance measures based on T.I.C, M.A.D, and E-VAF have been calculated satisfactorily. The statistical assessments for 100 independent trials using MWNN-GA-IPA in terms of minimum, S.I range, median, standard deviation, and maximum operatives further validate the worth and correctness of the proposed MWNN-GA-IPA. Furthermore, statistics analysis has been performed in the case of SIR dengue fever model based on dengue infection.

In future, the proposed MWNN-GA-IPA is proficient to solve the biological nonlinear systems, singular higher order nonlinear systems, and fluid dynamic systems.

Data Availability

This work is not based on any data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Side and M. S. M. Noorani, "A SIR model for spread of dengue fever disease (simulation for South Sulawesi, Indonesia and Selangor, Malaysia)," *World Journal of Modelling and Simulation*, vol. 9, no. 2, pp. 96–105, 2013.
- [2] S. Bhatt, P. W. Gething, O. J. Brady et al., "The global distribution and burden of dengue," *Nature*, vol. 496, no. 7446, pp. 504–507, 2013.
- [3] Y. C. Lai, Y. C. Chuang, C. C. Liu et al., "Antibodies against modified NS1 wing domain peptide protect against dengue virus infection," *Scientific reports*, vol. 7, no. 1, pp. 1–15, 2017.
- [4] F. Medina, J. F. Medina, C. Colón, E. Vergne, G. A. Santiago, and J. L. Muñoz-Jordán, "Dengue virus: isolation, propagation, quantification, and storage," *Current protocols in microbiology*, vol. 27, no. 1, 2012.
- [5] S. Cabrera-Romo, R. M. del Ángel, B. Recio-Tótoro et al., "Experimental inoculation of *Artibeus jamaicensis* bats with dengue virus serotypes 1 or 4 showed no evidence of sustained replication," *The American Journal of Tropical Medicine and Hygiene*, vol. 91, no. 6, pp. 1227–1234, 2014.
- [6] G. Li, P. Pan, Q. He et al., "Molecular epidemiology demonstrates that imported and local strains circulated during the 2014 dengue outbreak in Guangzhou, China," *Virologica Sinica*, vol. 32, no. 1, pp. 63–72, 2017.
- [7] J. Liu, Y. Liu, K. Nie et al., "Flavivirus NS1 protein in infected host sera enhances viral acquisition by mosquitoes," *Nature microbiology*, vol. 1, no. 9, pp. 1–11, 2016.
- [8] Y. G. Sánchez, Z. Sabir, and J. L. G. Guirao, "Design of a nonlinear SITR fractal model based on the dynamics of a novel coronavirus (COVID)," *Fractals*, vol. 28, no. 8, 2020.
- [9] L. Alphey, M. Benedict, R. Bellini et al., "Sterile-insect methods for control of mosquito-borne diseases: an analysis," *Vector Borne and Zoonotic Diseases*, vol. 10, no. 3, pp. 295–311, 2010.
- [10] A. H. Baumhover, A. J. Graham, B. A. Bitter et al., "Screw-worm control through release of sterilized Flies1," *Journal of Economic Entomology*, vol. 48, no. 4, pp. 462–466, 1955.
- [11] Z. Sabir, S. Saoud, M. A. ZahoorRaja, H. AbdulWahab, and A. Arbief, "Heuristic computing technique for numerical solutions of nonlinear fourth order Emden–Fowler equation," *Mathematics and Computers in Simulation*, vol. 178, 2020.
- [12] Z. Sabir, A. Fazli, P. Daniel, and J. L. G. Guirao, "Intelligence computing approach for solving second order system of Emden–Fowler model," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 1–16, 2020.
- [13] M. Umar, F. Amin, H. Wahab, and D. Baleanu, "Unsupervised constrained neural network modeling of boundary value corneal model for eye surgery," *Applied Soft Computing*, vol. 85, Article ID 105826, 2019.
- [14] Z. Sabir, M. A. Zahoor Raja, M. Umar, and M. Shoaib, "Neuro-swarm intelligent computing to solve the second-order singular functional differential model," *The European Physical Journal Plus*, vol. 135, no. 6, Article ID 474, 2020.
- [15] J. L. G. Guirao, Z. Sabir, and T. Saeed, "Design and numerical solutions of a novel third-order nonlinear emden–fowler

- delay differential model,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 7359242, 9 pages, 2020.
- [16] M. Umar, Z. Sabir, M. A. Zahoor Raja, and Y. G. Sánchezd, “A stochastic numerical computing heuristic of SIR nonlinear model based on dengue fever,” *Results in Physics*, vol. 19, Article ID 103585, 2020.
- [17] Y. Jiang, P. Wu, J. Zeng, Y. Zhang, Y. Zhang, and S. Wang, “Multi-parameter and multi-objective optimisation of articulated monorail vehicle system dynamics using genetic algorithm,” *Vehicle System Dynamics*, vol. 58, no. 1, pp. 74–91, 2020.
- [18] Z. Sabir, H. A. Wahab, M. Umar, M. G. Sakar, and M. A. Z. Raja, “Novel design of Morlet wavelet neural network for solving second order Lane-Emden equation,” *Mathematics and Computers in Simulation*, vol. 172, pp. 1–14, 2020.
- [19] Z. Tao, L. Huiling, W. Wenwen, and Y. Xia, “GA-SVM based feature selection and parameter optimization in hospitalization expense modeling,” *Applied Soft Computing*, vol. 75, pp. 323–332, 2019.
- [20] S. Sayed, M. Nassef, A. Badr, and I. Farag, “A nested genetic algorithm for feature selection in high-dimensional cancer microarray datasets,” *Expert Systems with Applications*, vol. 121, pp. 233–243, 2019.
- [21] D. Jude Hemanth and J. Anitha, “Modified genetic algorithm approaches for classification of abnormal magnetic resonance brain tumour images,” *Applied Soft Computing*, vol. 75, pp. 21–28, 2019.
- [22] M. A. Mohammed, M. K. Abd Ghani, R. I. Hamed, S. A. Mostafa, M. S. Ahmad, and D. A. Ibrahim, “Solving vehicle routing problem by using improved genetic algorithm for optimal solution,” *Journal of Computational Science*, vol. 21, pp. 255–262, 2017.
- [23] X. Luo, L. Niu, and S. Zhang, “An algorithm for traffic flow prediction based on improved sarima and GA,” *KSCE Journal of Civil Engineering*, vol. 22, no. 10, pp. 4107–4115, 2018.
- [24] M. Wilson, “Optimization of the radiation shielding capabilities of bismuth-borate glasses using the genetic algorithm,” *Materials Chemistry and Physics*, vol. 224, pp. 238–245, 2019.
- [25] D. Jahed Armaghani, M. Hasanipanah, A. Mahdiyar, M. Z. Abd Majid, H. Bakhshandeh Amnieh, and M. M. D. Tahir, “Airblast prediction through a hybrid genetic algorithm-ANN model,” *Neural Computing & Applications*, vol. 29, no. 9, pp. 619–629, 2018.
- [26] Y. Yang, B. Yang, S. Wang, F. Liu, Y. Wang, and X. Shu, “A dynamic ant-colony genetic algorithm for cloud service composition optimization,” *International Journal of Advanced Manufacturing Technology*, vol. 102, no. 1–4, pp. 355–368, 2019.
- [27] H. Motieghader, A. Najafi, B. Sadeghi, and A. Masoudi-Nejad, “A hybrid gene selection algorithm for microarray cancer classification using genetic algorithm and learning automata,” *Informatics in Medicine Unlocked*, vol. 9, pp. 246–254, 2017.
- [28] H. Zhang, J. Xie, J. Ge, Z. Zhang, and B. Zong, “A hybrid adaptively genetic algorithm for task scheduling problem in the phased array radar,” *European Journal of Operational Research*, vol. 272, no. 3, pp. 868–878, 2019.
- [29] M. Hassoon, M. S. Kouhi, M. Zomorodi-Moghadam, and M. Abdar, “Rule Optimization of Boosted C5.0 classification using genetic algorithm for liver disease prediction,” in *Proceedings of the 2017 International Conference on Computer and Applications (ICCA)*, September 2017.
- [30] C. Bertocchi, E. Chouzenoux, M. C. Corbineau, J. C. Pesquet, and M. Prato, “Deep unfolding of a proximal interior point method for image restoration,” *Inverse Problems*, vol. 36, no. 3, Article ID 34005, 2020.
- [31] A. Zanelli, A. Domahidi, J. Jerez, and M. Morari, “Forces nlp: an efficient implementation of interior-point methods for multistage nonlinear nonconvex programs,” *International Journal of Control*, vol. 93, no. 1, pp. 13–29, 2020.
- [32] J. Bleyer, “Advances in the simulation of viscoplastic fluid flows using interior-point methods,” *Computer Methods in Applied Mechanics and Engineering*, vol. 330, pp. 368–394, 2018.
- [33] D. Mangoni, A. Tasora, and R. Garziera, “A primal-dual predictor-corrector interior point method for non-smooth contact dynamics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 330, pp. 351–367, 2018.
- [34] N. P. Theodorakatos, “A nonlinear well-determined model for power system observability using Interior-Point methods,” *Measurement*, vol. 152, Article ID 107305, 2020.
- [35] F. Scott, P. Wilson, R. Conejeros, and V. S. Vassiliadis, “Simulation and optimization of dynamic flux balance analysis models using an interior point method reformulation,” *Computers & Chemical Engineering*, vol. 119, pp. 152–170, 2018.

Research Article

Application of DMAIC Cycle and Modeling as Tools for Health Technology Assessment in a University Hospital

Alfonso Maria Ponsiglione ¹, **Carlo Ricciardi** ², **Arianna Scala** ³, **Antonella Fiorillo** ²,
Alfonso Sorrentino ⁴, **Maria Triassi** ³, **Giovanni Dell'Aversana Orabona** ⁴,
and Giovanni Improta ³

¹Department of Electrical Engineering and Information Technology (DIETI), University of Naples "Federico II", Naples, Italy

²Department of Advanced Biomedical Sciences, University of Naples "Federico II", Naples, Italy

³Department of Public Health, University Hospital of Naples "Federico II", Naples, Italy

⁴Maxillofacial Surgery Unit, Department of Neurosciences, Reproductive and Odontostomatological Sciences, University Hospital of Naples "Federico II", Naples, Italy

Correspondence should be addressed to Arianna Scala; ariannascala7@gmail.com

Received 15 July 2020; Accepted 10 August 2021; Published 18 August 2021

Academic Editor: Daniel Espino

Copyright © 2021 Alfonso Maria Ponsiglione et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. The Health Technology Assessment (HTA) is used to evaluate health services, manage healthcare processes more efficiently, and compare medical technologies. The aim of this paper is to carry out an HTA study that compares two pharmacological therapies and provides the clinicians with two models to predict the length of hospital stay (LOS) of patients undergoing oral cavity cancer surgery on the bone tissue. **Methods.** The six Sigma method was used as a tool of HTA; it is a technique of quality management and process improvement that combines the use of statistics with a five-step procedure: "Define, Measure, Analyze, Improve, Control" referred to in the acronym DMAIC. Subsequently, multiple linear regression has been used to create two models. Two groups of patients were analyzed: 45 were treated with ceftriaxone while 48 were treated with the combination of cefazolin and clindamycin. **Results.** A reduction of the overall mean LOS of patients undergoing oral cavity cancer surgery on bone was observed of 40.9% in the group treated with ceftriaxone. Its reduction was observed in all the variables of the ceftriaxone group. The best results are obtained in younger patients (−54.1%) and in patients with low oral hygiene (−52.4%) treated. The regression results showed that the best LOS predictors for cefazolin/clindamycin are ASA score and flap while for ceftriaxone, in addition to these two, oral hygiene and lymphadenectomy are the best predictors. In addition, the adjusted R squared showed that the variables considered explain most of the variance of LOS. **Conclusion.** SS methodology, used as an HTA tool, allowed us to understand the performance of the antibiotics and provided variables that mostly influence postoperative LOS. The obtained models can improve the outcome of patients, reducing the postoperative LOS and the relative costs, consequently increasing patient safety, and improving the quality of care provided.

1. Introduction

Healthcare seeks to give improvements in the prevention, control, and treatment of diseases, but at the same time, it also deals with complications, inefficiencies, and other problems that put patients' safety at risk. Therefore, it is necessary to monitor the health services provided by applying management methods and tools to control quality [1].

Nowadays, several methodologies and approaches are used in healthcare to help in the clinical decision-making process [2–8], to aid physicians in defining the diagnosis and prognosis of patients [9–11], and to analyze quality improvement in hospital processes [12, 13]. A useful methodology for these purposes is the Health Technology Assessment (HTA), a multidisciplinary process for medical-clinical, social, organizational, economic, technological,

ethical, and legal implication analysis of health technology through the evaluation of efficiency, security, costs, and social and organizational impact [14, 15]. The technologies could be drugs, medical devices, vaccines, procedures, and, generally, all systems developed to solve a health problem and to improve the quality of life.

Parmar and Chan [16] used HTA methodology in urologic oncology. As a result of the rapid development of new cancer therapies, it is important to have a decision-making tool which leads to the choice of the right therapy in a short period of time. In this study, HTA was used as an approach that could help to guide value-based decision-making. An HTA model was developed for the evaluation of generic pharmaceutical products. This tool allows us to compare, both qualitatively and economically, equivalent drug preparation. HTA was employed to evaluate a new health technology for the thyroglobulin assay in patients with differentiated thyroid cancer. The authors used the Dynamic AHP as an HTA tool to reach the goal [17]; this paper proved also the utility of combining HTA with other managerial approaches.

Another promising tool to improve the quality of healthcare processes is Six Sigma (SS) [18–21]. Initially introduced in the manufacturing sector, today, it is widely developed in the health sector. SS relies on the “Define, Measure, Analyze, Improve, Control” cycle (DMAIC), which is a five-step procedure related to quality management and process improvement that exploits both statistical and managerial tools. Through this problem-solving strategy with a fixed structure, it is possible to analyze a process in order to improve its performance reducing the “natural variability” and carry out the “systematic control” of the critical variables to obtain a better result. The procedure is divided into the following phases: defining the project goals and customer (internal and external) requirements, measuring the process to determine current performance, analyzing and defining the root cause(s) of relevant defects, improving the process by eliminating defect root causes, and controlling future process performance. For the first time, Bill Smith developed this methodology in 1986 with the aim of reducing product or process defects that did not satisfy customers [18, 22]. DMAIC is then a framework used to enable the team to define and achieve set objectives [1, 23, 24].

From literature studies, it stands out the success that the strength of SS is founded not only in the manufacturing field but also in the health sector, where the SS DMAIC approach has been applied, for example, to improve first aid processes [25] and in the paramedical services [26]. Mahesh et al. [27] demonstrated how to reduce patients’ waiting time to receive a specialist medical visit at the Out-Patient Department of Cardiology in a private hospital in the city of Bangalore, and El-Eid et al. [28] have confirmed SS as an efficient and effective management tool to improve the patient discharge process, reducing patient discharge time. As well, other studies confirmed the validity of the methodology [13, 29–33], also in combination with other methods such as the Agile [34]. Ricciardi et al. [12] analyzed the introduction of the Diagnostic Therapeutic Assistance Path (DTAP),

employing Lean Thinking and SS methodology based on the DMAIC cycle. Furthermore, several studies show that the SS is often associated with Lean Thinking: this approach aims to improve services to meet customer needs by eliminating wastes and reducing costs [35–37]. The use of these methodologies has reported multiple benefits in healthcare; in fact, they have been used to improve clinical decision-making processes and to reduce the risk of healthcare-associated infections in surgery departments [38], while others have conducted studies to introduce prehospitalization to perform the necessary tests and examinations for hip and knee prosthetic surgery [29, 39].

The problem of healthcare infections is of great interest in many surgery departments, and it is an indicator of hospital efficiency, safety, and quality. Scotton et al. [40] conducted a study whose purpose was to analyze infections in patients after Salvage Laryngectomy (SL) and review the potential impact of the antibiotic prophylaxis adopted. The results showed that infection rates after SL were high, and univariate analysis demonstrated risk variables that had a significant correlation with infection, so the antibiotic regimen is probably ineffective. Other authors [41–48] presented an overview of current evidence-based best practices in the use of prophylactic antibiotics in head and neck cancer surgery; indeed, this type of patient is at high risk of developing complications after surgery. Thus, they reported that prophylactic antibiotics helped significantly reduce the risk of infection [49]. However, short four-dose antibiotic regimens for 24 hours are as effective as prolonged cycles, regardless of the complexity of the procedure [50–53]. In the same framework, the research of Egan et al. [54] discusses the use of the SS focusing on therapy with antimicrobial gentamicin, which requires good practice in selecting the dose and monitoring serum levels. They found a new dosage with a standardized sampling, a monitoring program, and a new timing of drug delivery that maximized local capacities. In light of the above-mentioned studies, it emerges the importance of choosing correct prophylactic antibiotics to manage patients appropriately after surgical interventions.

To this aim, in our recent study [55], SS was employed to compare the use of antibiotics in patients undergoing oral cancer surgery on bone tissue. Starting from the previous promising results, in this work, two antibiotics, ceftriaxone and the combination of cefazolin and clindamycin, are compared in order to understand which one reduces the postoperative length of hospital stay (LOS) for patients undergoing oral cavity cancer surgery on the bone tissue. In this study, it is taken into consideration the clinical factor because the two antibiotics are quite similar from a safety, legal, ethical, economic, and technological point of view. Six Sigma (SS) methodology is applied as a tool of HTA in order to achieve the aim. SS was used to analyze the influence of some clinical variables (ASA score, age, gender, oral hygiene, diabetes, and cardiovascular diseases) on the Critical to Quality (CTQ) (postoperative LOS). Patients’ postoperative LOS can be described as the duration of time after a patient’s surgery until the day of discharge.

The novelty of this new study is the use of the DMAIC cycle as an HTA tool including a modeling phase. This would enable healthcare providers to understand the performance of antibiotics, improving patients' outcomes, reducing postoperative LOS and related costs, consequently, increasing patient safety, and improving the quality of care provided. After applying DMAIC, a modeling study was conducted through a multinomial linear regression; in particular, it was applied to obtain two models capable of predicting postoperative LOS for each antibiotic. In order to do this, we included the surgical variables that were considered in the previous study [55].

2. Materials and Statistical Tools

SS and subsequently the modeling phase were used to implement the HTA methodology. In detail, deploying the DMAIC cycle, characteristic of SS, means developing five phases:

- (1) The Define phase identifies the customers and the objectives to be reached will be established [27] allowing a team to identify the problem
- (2) The Measure phase defines the main characteristics of the process and the parameters that will lead to improvement [56]
- (3) The Analyze phase is used to understand the influence of the collected variables on the CTQ or to evaluate the data collected in the previous phases of the study using various analytical tools available such as regression analysis, fishbone diagram, tree diagrams, and brainstorming
- (4) The Improve phase employs all the previous analyses to design changes in a process and to improve the performance, i.e., introducing a new antibiotic protocol
- (5) The Control phase is employed to monitor the whole process and, in this research, to compare the performance of the drugs

SS led the way for the development of the modeling phase, providing us with information about all the variables. Modeling allowed us to enrich the univariate analysis with a multivariate one and to implement a tool able to predict the postoperative LOS for each patient. These models will be very useful for both ward management and hospital management. Predicting the LOS of a patient determines a more efficient hospital bed organization, a better management of nurses and doctors on duty, and lastly, a cost reduction for hospitals. Thus, combining SS and modeling could be considered a valuable tool for HTA methodology.

In conclusion, the purpose of this paper is to assess the performance of two antibiotics, cefazolin plus clindamycin [57, 58] and ceftriaxone [59], through an HTA by using SS and modeling as a tool in the framework of oral cavity cancer surgery on bone tissues.

2.1. The Clinical Case Study. In this study, two groups of patients with oral cancer starting from the bone were analyzed: the first one was treated with ceftriaxone between 2006 and 2011, while the second one was treated with cefazolin and

clindamycin between 2011 and 2019. The cefazolin group consisted of 54 patients, while the other by 51 patients. Oral cancer is the sixth most common cancer in the world [60] but the ones starting from the jaws are rare. The majority of the oral cancers affecting the bone derives from the epithelial quote of the oral mucosa, but there are also cancers that originally start from the bones, which are rare. Sarcomas are very rare tumors in the head and neck district, osteosarcoma being the most common of them [61]. They represent 1% of all the malignancies affecting the head and neck [62]. The incidence of sarcomas starting from the mandibles ranges from 4% to 10% [63]. In this study, we decided to analyze also those patients affected by ameloblastomas, which is not actually a malignant neoplasm. This choice is due to the fact that in the case of big ameloblastomas affecting the jaws, a big removal of tissue and reconstruction with the same surgical techniques used for patients affected by oral bone cancers are often required. The data was taken from printed medical records. Statistical tests, useful for analyses, were carried out with IBM SPSS.

For the collection of data, some inclusion and exclusion criteria were taken into consideration:

- (i) All patients were included without exclusion due to medical history (gender, age, cardiovascular diseases, diabetes, oral hygiene, American Society of Anaesthesiologists (ASA) Score)
- (ii) Patients with cancers starting from the bones or starting from the oral mucosa and then affecting the bone were included. We also included patients with ameloblastomas because of their osteolytic patterns
- (iii) Patients treated in "day surgery" were excluded
- (iv) Patients with too many missing data were not included because they would compromise the analysis
- (v) Patients with a change of the antibiotic therapy during their recovery, because no evidence of efficacy, were not included in the analysis, but their number was recorded as it is a qualitative indicator of treatment failure
- (vi) Patients allergic to cefazolin and clindamycin or ceftriaxone were excluded

As regards the Unit of Maxillofacial Surgery, the ward consists of 9 rooms with 22 beds for the patients and some more rooms for surgeons and nurses. The Operatory Block of the Department disposes of two operating rooms.

Oncological maxillofacial surgery is a branch of maxillofacial surgery which deals with the surgical approach to head and neck malignancies and the reconstruction of the lost tissues [64].

When no allergy was described, from 2006 to 2011, a postoperative antibiotic protocol with ceftriaxone was used. Since 2011, there has been a shift to the use of the association of cefazolin plus clindamycin as postoperative antibiotic prophylaxis.

2.2. The Development of the Six Sigma: The Define Phase. The purpose of the "Define" phase is to define a multidisciplinary workgroup and to divide the tasks for the analysis.

The team consists of clinicians from the Maxillofacial Department of the University Hospital “Federico II” of Naples, an economist, and biomedical engineers with experience in health management. The team was responsible for collecting and analyzing data of patients with oral cavity cancer considering the influence of some variables. The sample and the leader supervised and coordinated the study and interpretation of the data. A project diagram was created to define the problem to be solved:

- (i) *Project Title.* Health Technology Assessment between two antibiotics in the context of Maxillofacial Surgery
- (ii) *Question.* Investigation of the best antibiotics in the analyzed context
- (iii) *Critical to Quality.* Postoperative LOS
- (iv) *Target.* Realize corrective measures to reduce the CTQs
- (v) *Deliverables.* The performance of cefazolin/clindamycin and ceftriaxone, the outcome of patients, reducing postoperative LOS, and the related costs
- (vi) *Timeline:*
 - (1) Define: January 2010
 - (2) Measure: January 2010
 - (3) Analyze: January 2010
 - (4) Improve: January 2011
 - (5) Control: 2011–2018
- (vii) *In Scope.* Oral cavity cancer surgery on bone tissues. Maxillofacial surgery in the University Hospital of Naples “Federico II”
- (viii) *Out of Scope.* All the other structures and interventions and drugs
- (ix) *Financial.* No funding to reach the target
- (x) *Business Need.* Identifying the best antibiotic for the surgery under examination

2.3. Dataset Description: The Measure Phase. The data collected from the medical records at the Department of Maxillofacial Surgery were selected according to the inclusion and exclusion criteria. After applying the inclusion and exclusion criteria, the first sample of data concerned patients treated with ceftriaxone from 2006 to 2011 (45 patients), and the other sample of data (48 patients) was referred to patients treated with cefazolin and clindamycin from 2011 to 2019. The variables used to compare the two antibiotics were

- (i) Gender
- (ii) Age
- (iii) American Society of Anaesthesiologists (ASA) Score
- (iv) Quality of oral hygiene
- (v) Diabetes
- (vi) Cardiovascular diseases

Other variables were analyzed through univariate analysis in a previous study [55]; thus, they were included

only in the modeling phase. Descriptive characteristics of the dataset were carried out for the postoperative LOS variables: the results for cefazolin/clindamycin were, respectively, an average of 16.51 days and a variance of 62.21. Instead, the results for ceftriaxone were an average of 9.75 days and a variance of 66.81.

We drew a histogram (Figure 1) showing the mean postoperative LOS of patients, measured in days, submitted to the administration of cefazolin/clindamycin according to each variable. The highest average LOS is for patients with a high ASA score, while the lowest is for patients with a low ASA score.

Figure 2 shows the distribution of mean postoperative LOS of patients who used ceftriaxone. Patients below the age of 51 have the highest mean LOS, whereas those without cardiovascular disease have the lowest mean LOS.

2.4. Statistical Analysis: The Analyze Phase. In Figure 3, patients’ pathway is shown from the arrival at the hospital to the discharge. They arrived at the hospital; then, if they receive a previous prehospitalization, they undergo surgery directly; otherwise, they are subjected to preoperative activities before surgery. Finally, if there are complications after the surgery, the patient undergoes postoperative activities; otherwise, they will be discharged after fewer days.

A Kolmogorov–Smirnov test showed a p value lower than 0.0001. In order to understand the variables that could influence the postoperative LOS in the ceftriaxone group, nonparametric tests were employed: Mann–Whitney and Kruskal–Wallis (only for age). In this case, some significant p values were found for age and ASA score while the p value of cardiovascular disease was almost significant (p value = 0.066) (Table 1).

A box diagram was developed and is shown in Figure 4, which clearly highlights the decrease in the ceftriaxone group of LOS, measured in days.

The Control phase allowed us to monitor and guarantee the sustainability of the long-term continuous improvement of the performance. Thus, the team identified the following actions:

- (i) Periodic review meetings to evaluate the maxillofacial surgery process
- (ii) Internal audit to verify the performance of antibiotics
- (iii) Production of reports that highlight the trend of patients’ postoperative patients measured in days

After analyzing the data according to the DMAIC cycle, the modeling phase started by implementing the multiple linear regression. It is also known simply as multiple regression and is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression is to model the linear relationship between the explanatory (independent) variables and response (dependent) variables. In other words, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

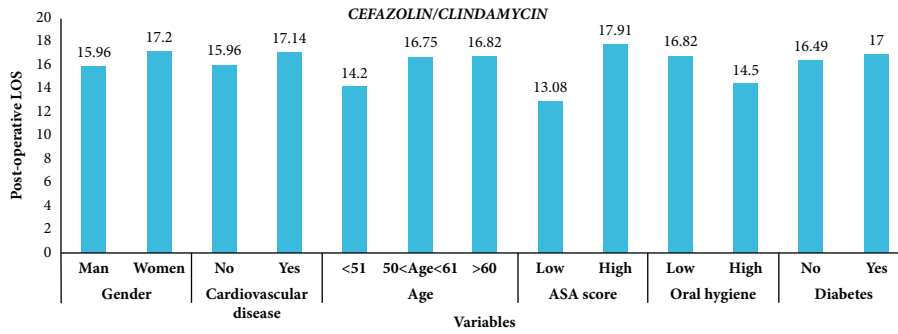


FIGURE 1: Mean postoperative LOS for each mode of variables regarding cefazolin/clindamycin.

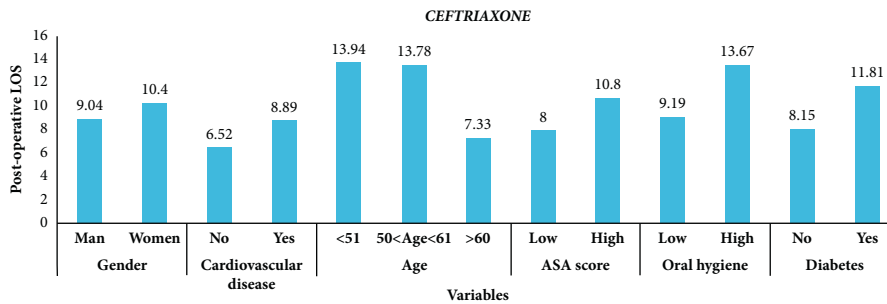


FIGURE 2: Mean postoperative LOS for each mode of variables regarding ceftriaxone.

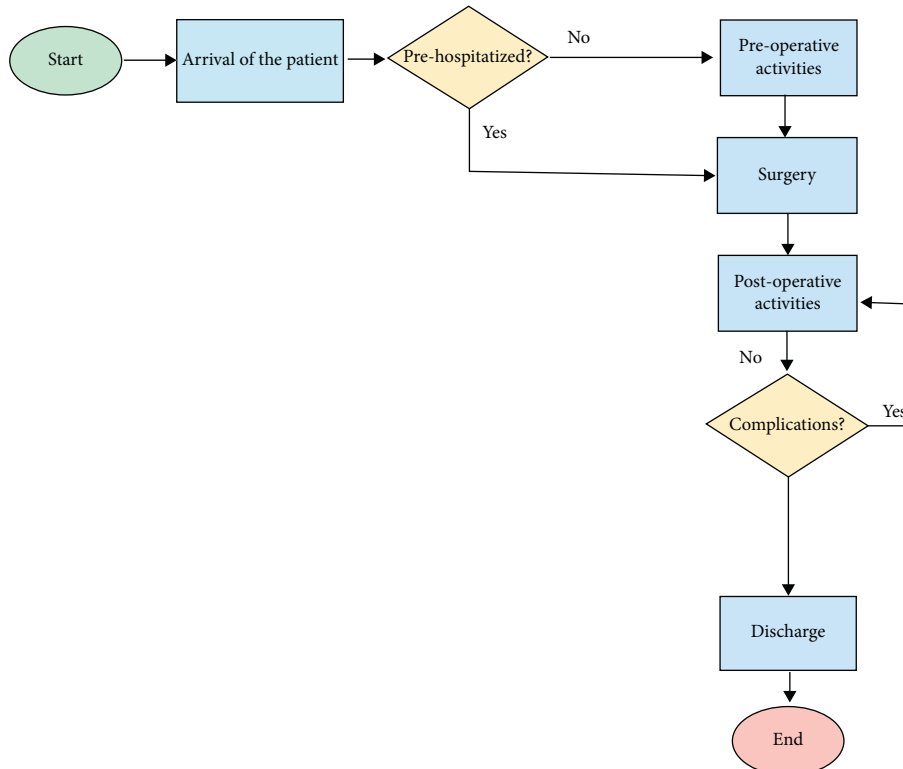


FIGURE 3: The flowchart of the hospitalization process for patients undergoing oncologic surgery at the Maxillofacial Department of the University Hospital of Naples “Federico II.”

In this study, it was used to obtain a model capable of predicting the postoperative LOS for each patient undergoing oral cavity cancer surgery on the bone. In order to

obtain the best models, we considered also the surgical variables that were studied in a previous research on the same topic [55]. Therefore, the considered variables in order

TABLE 1: The analysis of potential factors influencing postoperative LOS for the “ceftriaxone” group.

Variable	Category	N	LOS (mean \pm std. dev.)	p value
Gender	Men	25	9.04 \pm 7.49	0.669
	Women	23	10.40 \pm 9.02	
Age	<51	21	6.52 \pm 5.33	0.013*
	50 < age < 61	9	8.89 \pm 6.92	
	>60	18	13.94 \pm 10.04	
ASA score	Low	30	7.33 \pm 5.84	0.007
	High	18	13.78 \pm 10.15	
Oral hygiene	Low	30	8.00 \pm 6.74	0.306
	High	18	10.80 \pm 9.00	
Diabetes	No	42	9.19 \pm 8.05	0.213
	Yes	6	13.67 \pm 9.46	
Cardiovascular disease	No	27	8.15 \pm 7.48	0.066
	Yes	21	11.81 \pm 8.92	

*Kruskal–Wallis test.

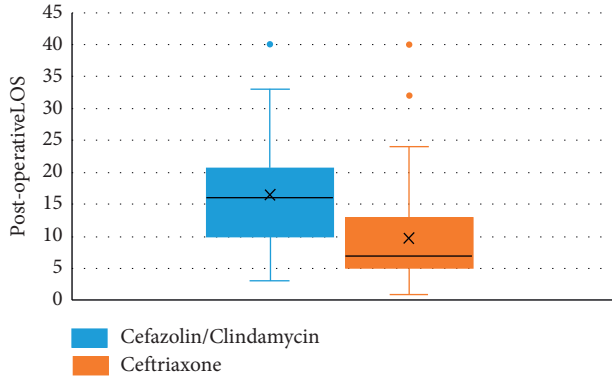


FIGURE 4: Boxplot of the mean postoperative LOS for “cefazolin/clindamycin” and “ceftriaxone” groups.

to implement the model were 11: gender, age, ASA score, the quality of oral hygiene, diabetes, cardiovascular diseases, tracheotomy, lymphadenectomy, infections, dehiscence, and flap.

3. Results

3.1. Statistical Analysis for Cefazolin plus Clindamycin. The Kolmogorov–Smirnov test was applied to investigate the distribution of the postoperative LOS data regarding cefazolin/clindamycin; a p value of 0.200 indicated a normality distribution. Thus, to investigate the variables potentially influencing postoperative LOS, t -test and ANOVA were employed. The results are represented in Table 2. No significance was found in the tests, but the difference between postoperative LOS in each category gave insights about a potential influence in many of the variables; the ASA score was almost significant.

3.2. Comparison between the Two Antibiotics: The Control Phase. The Kolmogorov–Smirnov test showed a p value of less than 0.0001; i.e., the data were not normally distributed.

TABLE 2: The analysis of potential factors influencing postoperative LOS for the “cefazolin/clindamycin” group.

Variable	Category	N	LOS (mean \pm std. dev.)	p value
Gender	Men	25	15.96 \pm 7.32	0.606
	Women	20	17.20 \pm 8.68	
Age	<51	5	14.20 \pm 7.26	0.793*
	50 < age < 61	12	16.75 \pm 9.11	
	>60	28	16.82 \pm 7.65	
ASA score	Low	13	13.08 \pm 6.69	0.062
	High	32	17.91 \pm 8.00	
Oral hygiene	Low	39	16.82 \pm 8.17	0.509
	High	6	14.50 \pm 5.89	
Diabetes	No	43	16.49 \pm 8.07	0.930
	Yes	2	17.00 \pm 0.00	
Cardiovascular disease	No	24	15.96 \pm 8.65	0.621
	Yes	21	17.14 \pm 7.07	

*ANOVA test.

The results of the comparison between the two antibiotics through Mann–Whitney and Kruskal–Wallis tests with an alpha level of 0.05 are shown in Table 3. Overall, the difference in postoperative LOS between the cefazolin/clindamycin and ceftriaxone groups was statistically significant with a reduction of 40.9%. All tests were statistically significant among the mode of variables, except for older patients (>60 years with a p value of 0.117). The greatest reduction in postoperative LOS results in younger patients (<51 years with a reduction of 54.1%) and people with low oral hygiene (52.4%).

Table 4 shows the results of a study regarding the frequencies of each variable, obtained by performing a chi-square test. A statistically significant difference between the occurrences of cefazolin/clindamycin and ceftriaxone groups was obtained according to age, ASA score, and oral hygiene.

3.3. Combining SS and Modeling. The statistical analysis was useful for the subsequent modeling phase. As mentioned in the introduction, in this phase, we also considered some surgical variables analyzed in a preceding paper [55]. For both antibiotic protocols, the multiple linear regression was implemented obtaining two predictive models whose equations are shown as follows:

$$y_1 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \varepsilon_1, \quad (1)$$

$$y_2 = \partial_1 x_1 + \partial_2 x_2 + \partial_3 x_3 + \partial_4 x_4 + \partial_5 x_5 + \partial_6 x_6 + \varepsilon_2, \quad (2)$$

where y_1 represents the LOS of patients treated with cefazolin/clindamycin, y_2 the LOS of patients treated with ceftriaxone, x_i the considered variables, β_i and ∂_i the regression coefficients, and ε_i the errors.

Before carrying out the regression analysis, it is necessary to verify, for both antibiotics, the hypotheses given in Table 5 which also contains references to the additional material provided in order to give more details on these verifications.

TABLE 3: The complete comparative statistical analysis. Mann–Whitney and Kruskal–Wallis were used, respectively, for dichotomous groups and for the age group.

Variable	Category	Cefazolin/clindamycin (mean \pm std. dev.)	Ceftriaxone (mean \pm std. dev.)	Difference (%)	<i>p</i> value
All patients		16.51 \pm 7.89	9.75 \pm 8.26	–40.9	<0.0001
Gender	Men	15.96 \pm 7.32	9.04 \pm 7.49	–43.4	0.003
	Women	17.20 \pm 8.68	10.40 \pm 9.02	–39.5	0.002
Age	<51	14.20 \pm 7.26	6.52 \pm 5.33	–54.1	0.015*
	50 < age < 61	16.75 \pm 9.11	8.89 \pm 6.92	–46.9	0.028*
	>60	16.82 \pm 7.65	13.94 \pm 10.04	–17.4	0.117*
ASA score	Low	13.08 \pm 6.69	7.33 \pm 5.84	–44.0	0.007
	High	17.91 \pm 8.00	13.78 \pm 10.15	–23.1	0.042
Oral hygiene	Low	16.82 \pm 8.17	8.00 \pm 6.74	–52.4	0.001
	High	14.50 \pm 5.89	10.80 \pm 9.00	–25.5	0.040
Diabetes	No	16.49 \pm 8.07	9.19 \pm 8.05	–44.2	<0.0001
	Yes	17.00 \pm 0.00	13.67 \pm 9.46	n.a.	n.a.
Cardiovascular disease	No	15.96 \pm 8.65	8.15 \pm 7.48	–48.9	<0.0001
	Yes	17.14 \pm 7.07	11.81 \pm 8.92	–31.2	0.012

*Kruskal–Wallis test; n.a.: not applicable.

TABLE 4: The analysis of the frequencies for each variable is performed through a chi-square test.

Variable	Category	Cefazolin/clindamycin (N)	Ceftriaxone (N)	<i>p</i> value
Gender	Men	25	25	0.737
	Women	20	23	
Age	<51	5	21	0.002
	50 < age < 61	12	9	
	>60	28	18	
ASA score	Low	13	30	0.001
	High	32	18	
Oral hygiene	Low	39	30	0.008
	High	6	18	
Diabetes	No	43	42	0.166
	Yes	2	6	
Cardiovascular disease	No	24	27	0.778
	Yes	21	21	

TABLE 5: Verification of the assumptions of multiple regression models for both antibiotics and reference to corresponding Supplementary Material items.

Assumption	Description	Reference to Supplementary Material
Linearity	Verify if a linear relationship exists between the dependent variable and each predictor of the model	Figures S1 and S2
Independence of residuals	Verify if the errors of the model are independent	Tables S1
Collinearity	Verify if the predictors are not linearly correlated with each other	Table S2
Outliers	Verify if there are influential cases biasing the model	Figure S3
Normality of the residuals	Verify if the errors of the model are normally distributed	Figure S4
Homoscedasticity	Verify if the variance of the errors of the model is constant	Figure S5

As shown in equations (1) and (2), not all variables were considered for both models. In particular, 8 variables were included for cefazolin/clindamycin (ASA score, diabetes, cardiovascular disease, tracheotomy, lymphadenectomy, infections, dehiscence, and flap) while 6 variables were included for ceftriaxone (ASA score, oral hygiene, diabetes, cardiovascular disease, lymphadenectomy, and flap). The exclusion criteria of variables in each model were as follows:

- (i) Gender and age were excluded in order to obtain models based on clinical factors
- (ii) Oral hygiene was excluded from the cefazolin/clindamycin model because it did not respect the “absence of multicollinearity” hypothesis; i.e., there was a dependency between it and the ASA score variable. Since ASA score had a lower *p* value in the previous analyses of DMAIC than oral hygiene, the latter was excluded

TABLE 6: Regression coefficients, errors, and p value for cefazolin/clindamycin model.

Variables	Unstandardized regression coefficients (cefazolin/clindamycin)		
	Regression coefficients (β_i)	Std. error	p value
ASA score	3.406	0.506	0.000
Diabetes	1.025	4.066	0.803
Cardiovascular disease	2.541	1.707	0.147
Tracheotomy	0.022	2.366	0.993
Lymphadenectomy	2.816	2.139	0.198
Infections	2.790	3.383	0.416
Dehiscence	2.636	2.507	0.301
Flap	3.617	1.824	0.056

TABLE 7: Regression coefficients, errors, and p value for ceftriaxone model.

Variables	Unstandardized regression coefficients (ceftriaxone)		
	Regression coefficients (δ_i)	Std. error	p value
ASA score	2.272	0.609	0.001
Oral hygiene	0.873	0.358	0.020
Diabetes	4.938	2.546	0.600
Cardiovascular disease	-0.423	1.732	0.808
Lymphadenectomy	14.174	5.592	0.015
Flap	6.991	2.340	0.005

TABLE 8: Coefficient of determination, adjusted R squared, and standard errors of the two models.

	Cefazolin/clindamycin	Ceftriaxone
R^2	0.914	0.847
Adjusted R squared	0.892	0.823
Std. error	5.218	4.799

- (iii) Infections and dehiscence were excluded from the ceftriaxone model because no patient has experienced them. Similarly, the tracheotomy variable was excluded because there was only one case and it was not enough

Tables 6 and 7 show the regression coefficients, errors, and statistical significance obtained for each variable.

The results show that for cefazolin/clindamycin the ASA score is statistically significant and the flap is very close to the p value of 0.05. Similarly, for ceftriaxone the ASA score and the flap are variables that have a significant effect on LOS, as well as oral hygiene and lymphadenectomy.

A summary of the two models is given in Table 8. In particular, there are the coefficient of determination (R^2), the adjusted R squared, and the standard error of the estimate.

Since the two models have a different number of predictors, in addition to the R^2 , the adjusted R squared has also been reported; it is a modified version of R^2 , adjusted

according to the number of predictors in the model. Although there are also other variables affecting LOS, the results obtained indicate that, for both antibiotics, about 82–89 percent of the variance in LOS is explained by the selected variables.

4. Discussion and Conclusion

Over the past few years, the healthcare sector has paid attention to cost increases, mainly due to the drop of refunds, and to improve the experience of patients. In this scenario, the HTA provides health leaders with a useful tool to improve the efficiency and effectiveness of clinical processes; this tool has become fundamental in healthcare due to the high amount of medical device patents that have been required in the last decades [65]. In the literature, some studies applied the HTA to support decision-making processes regarding the purchase of medical devices [66] or drug refund policies [67, 68], while only a few works present an application of the HTA for evaluating the introduction of new antibiotic prophylaxis. In this study, we tackled this issue by employing a combination of both SS and HTA. In particular, encouraged by the results achieved in previously published studies [7, 55], here we adapted the framework of the SS DMAIC cycle to build a tool that could support the HTA of a new antibiotic prophylaxis procedure for patients undergoing oral cancer surgery of the bone. The assessment has been made taking into account a healthcare key performance indicator, which is the postoperative LOS. Indeed, the LOS is a useful metric to determine the economic, organizational, and clinical impact of healthcare services. In this work, a multiple regression model has been integrated within the SS framework to investigate the relationships between a prolonged LOS and the prophylaxis procedure in order to determine the impact of the introduction of a new antibiotic on the hospital stay. When framed into the Improve phase of the SS DMAIC cycle, the regression model helped in determining the effect of the new antibiotic prophylaxis on the postoperative LOS and enabled a comparison between the two antibiotics, thus providing an additional informative tool to support the decision-making process, in accordance with our previous works [7, 55]. The results obtained from the comparative statistical analysis (Table 3) showed a 41% reduction in the LOS for patients treated with ceftriaxone compared to those treated with cefazolin/clindamycin, with the highest decrease achieved among younger patients (–54.1%). This could be due to the better response of younger patients toward the performed surgical procedure, as opposed to older patients, whose surgical intervention can be influenced by possible comorbidities and other variables, in accordance with the literature [69, 70]. The modeling phase with the two regression models (Tables 6 and 7) enabled the identification of the variables, among demographic, clinical, and surgical ones as considered in a previous study [55], which influence the postoperative LOS the most and provided promising tools for the prediction of the LOS in patients undergoing oral cavity cancer surgery on the bone who are treated with cefazolin/clindamycin or with ceftriaxone. Of note, during the whole study's range of time, the choice of the antibiotics was completely independent

of the research. Indeed, the antibiotic to be administered was defined by the hospital's protocols which change the antibiotic choice in 2011 according to the new trends of therapy described in the medical literature.

In summary, the proposed approach confirmed the value of combining both the SS DMAIC approach and modeling, which can serve as a tool to support HTA processes for understanding the optimal therapeutic approach.

In conclusion, this HTA study confirmed and further extended the results achieved and presented in the literature which considered the ceftriaxone as the best option for patients undergoing oral cancer surgery on bone tissue [55] and provides the health policy with two important results: the antibiotic which reduces the postoperative LOS and two models which predict it. Succeeding in predicting the postoperative LOS of a patient could lead to many benefits for both the hospital and patients. Indeed, the hospital could better manage all its resources, reduce waste and costs, and improve the understanding of patients' needs, which are all aims of an SS project; meanwhile, the patients could experience a better quality of care and a lower LOS.

The evaluation of antibiotic performance is an important topic, as it is linked to healthcare-associated infections in hospitals, as evidenced by studies in the literature. This paper evaluates the performance of antibiotics considering the most important variables in the maxillofacial area. In addition, the DMAIC approach implies a positive advantage, giving support to the medical staff in the decision-making process of antibiotic administration, reducing the gap between practice and theory. Therefore, the reduction of postoperative LOS and the rate of infections of patients undergoing oral cavity cancer surgery benefit both the hospital and patients: patients satisfied in terms of a few days of hospitalization and effective and efficient therapy, while the hospital has more available beds and saves costs of managing patients with complications.

Abbreviations

ASA:	American Society of Anaesthesiologists
CTQ:	Critical to quality
DMAIC:	Define, Measure, Analyze, Improve, Control
LOS:	Length of hospital stay
SPSS:	Statistical Software for the Social Sciences
SS:	Six Sigma.

Data Availability

Data are not present in a publicly accessible repository. Data could be made available upon reasonable request to the authors.

Conflicts of Interest

The authors declare they have no conflicts of interest.

Authors' Contributions

Alfonso Maria Ponsiglione, Carlo Ricciardi, Arianna Scala, Giovanni Dell'Aversana Orabona and Giovanni Improta contributed equally to this work.

Supplementary Materials

Details on the multiple regression model assumptions check (as briefly summarized in Table 5 of the manuscript) are reported in the attached Supplementary Material file. (*Supplementary Materials*)

References

- [1] S. Akifuddin and F. Khatoun, "Reduction of complications of local anaesthesia in dental healthcare setups by application of the six sigma methodology: a statistical quality improvement technique," *Journal of Clinical and Diagnostic Research: Journal of Clinical and Diagnostic Research*, vol. 9, no. 12, pp. ZC34–38, 2015.
- [2] S. Domínguez and M. C. Carnero, "Fuzzy multicriteria modelling of decision making in the renewal of healthcare technologies," *Mathematics*, vol. 8, no. 6, 2020.
- [3] C. Ricciardi, A. M. Ponsiglione, G. Converso, I. Santalucia, M. Triassi, and G. Improta, "Implementation and validation of a new method to model voluntary departures from emergency departments," *Mathematical Biosciences and Engineering*, vol. 18, no. 1, pp. 253–273, 2021.
- [4] A. Scala, A. M. Ponsiglione, I. Loperto et al., "Lean six sigma approach for reducing length of hospital stay for patients with femur fracture in a university hospital," *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, 2021.
- [5] G. Improta, A. M. Ponsiglione, G. Parente et al., "Evaluation of medical training courses satisfaction: qualitative analysis and analytic hierarchy process," in *Proceedings of the 8th European Medical and Biological Engineering Conference*, T. Jarm, A. Cvetkoska, S. K. Mahnič, and D. Miklavcic, Eds., Springer International Publishing, Portorož, Slovenia, pp. 518–526, November 2020.
- [6] M. J. Glover, E. Jones, K. L. Masconi, M. J. Sweeting, and S. G. Thompson, "Discrete event simulation for decision modeling in health care: lessons from abdominal aortic aneurysm screening," *Medical Decision Making*, vol. 38, no. 4, pp. 439–451, 2018.
- [7] A. M. Ponsiglione, C. Ricciardi, G. Improta et al., "A Six Sigma DMAIC methodology as a support tool for Health Technology Assessment of two antibiotics," *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 3469–3490, 2021.
- [8] A. Glaize, A. Duenas, C. D. Martinelly, and I. Fagnot, "Healthcare decision-making applications using multicriteria decision analysis: a scoping review," *Journal of Multi-Criteria Decision Analysis*, vol. 26, no. 1–2, pp. 62–83, 2019.
- [9] G. D'Addio, C. Ricciardi, G. Improta, P. Bifulco, and M. Cesarelli, "Feasibility of machine learning in predicting features related to congenital nystagmus," in *Proceedings of the Mediterranean Conference on Medical and Biological Engineering and Computing*, Springer, Coimbra, Portugal, September 2019.
- [10] C. Ricciardi, V. Cantoni, G. Improta et al., "Application of data mining in a cohort of Italian subjects undergoing myocardial perfusion imaging at an academic medical center," *Computer Methods and Programs in Biomedicine*, vol. 189, 2020.
- [11] C. Ricciardi, K. J. Edmunds, M. Recenti et al., "Assessing cardiovascular risks from a mid-thigh CT image: a tree-based machine learning approach using radiodensitometric distributions," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020.

- [12] C. Ricciardi, A. Fiorillo, A. S. Valente et al., "Lean Six Sigma approach to reduce LOS through a diagnostic-therapeutic-assistance path at A," *O.R.N.A. Cardarelli. The TQM Journal*, vol. 31, no. 5, pp. 657–672, 2019.
- [13] C. Ricciardi, G. Balato, M. Romano, I. Santalucia, M. Cesarelli, and G. Improta, "Fast track surgery for knee replacement surgery: a lean six sigma approach," *TQM Journal*, vol. 32, no. 3, 2020.
- [14] R. N. Battista and M. J. Hodge, "The evolving paradigm of health technology assessment: reflections for the millennium," *Canadian Medical Association Journal: Canadian Medical Association Journal*, vol. 160, no. 10, pp. 1464–1467, 1999.
- [15] C. Favaretti, A. Cicchetti, G. Guarrera, M. Marchetti, and W. Ricciardi, "Health technology assessment in Italy," *International Journal of Technology Assessment in Health Care*, vol. 25, no. Suppl 1, pp. 127–133, 2009.
- [16] A. Parmar and K. K. W. Chan, "Health technology assessment methodology in metastatic renal cell carcinoma," *Nature Reviews Urology*, vol. 17, no. 1, pp. 3–5, 2020.
- [17] G. Improta, A. Perrone, M. A. Russo, and M. Triassi, "Health technology assessment (HTA) of optoelectronic biosensors for oncology by analytic hierarchy process (AHP) and Likert scale," *BMC Medical Research Methodology*, vol. 19, no. 1, 2019.
- [18] J. Antony, P. Palsuk, S. Gupta, D. Mishra, and P. Barach, "Six Sigma in healthcare: a systematic review of the literature," *International Journal of Quality & Reliability Management*, vol. 35, no. 5, pp. 1075–1092, 2018.
- [19] R. M. Alhamali, "Success factors and benefits of six sigma implementation in hospitals: a systematic review," *Business and Management Studies*, vol. 5, no. 3, pp. 1–10, 2019.
- [20] T. T. Allen, S.-H. Tseng, K. Swanson, and M. A. McClay, "Improving the hospital discharge process with six sigma methods," *Quality Engineering*, vol. 22, no. 1, pp. 13–20, 2009.
- [21] M. Arafeh, M. A. Barghash, N. Haddad et al., "Using six sigma DMAIC methodology and discrete event simulation to reduce patient discharge time in king hussein cancer center," *Journal of Healthcare Engineering*, vol. 2018, Article ID 3832151, 18 pages, 2018.
- [22] V. Sawalakhe, S. V. Deshmukh, and R. R. Lakhe, *Evaluating Performance of Testing Laboratory Using Six Sigma 1 Pranil*, 2016.
- [23] M. A. O. Barrios and H. F. Jiménez, "Use of six sigma methodology to reduce appointment lead-time in obstetrics outpatient department," *Journal of Medical Systems*, vol. 40, no. 10, p. 220, 2016.
- [24] A. George, A. M. Joseph, S. Kolencherry et al., "Application of six sigma dmaic methodology to reduce medication errors in a major trauma care centre in india," *Indian Journal Of Pharmacy Practice*, vol. 11, no. 4, pp. 182–187, 2018.
- [25] M. Fieri, N. F. Ranney, E. B. Schroeder, E. M. A. Van, and A. H. Stone, "Analysis and improvement of patient turnaround time in an Emergency Department," in *Proceedings of the 2010 IEEE Systems and Information Engineering Design Symposium*, pp. 239–244, Charlottesville, VA, USA, April 2010.
- [26] A. A. Baddour and H. A. Saleh, "Use six sigma approach to improve healthcare workers safety," *International Journal of Pure and Applied Sciences and Technology*, vol. 18, no. 1, 2013.
- [27] B. P. Mahesh, B. Soragaon, and A. R. Annigeri, "Reduction of patient wait time at a multi-specialty hospital using DMAIC methodology and factor Analysis," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 309–312, 2018.
- [28] G. R. E. El-Eid, R. Kaddoum, H. Tamim, and E. A. Hitti, "Improving hospital discharge time: a successful implementation of six sigma methodology," *Medicine*, vol. 94, no. 12, 2015.
- [29] G. Improta, G. Balato, M. Romano et al., "Improving performances of the knee replacement surgery process by applying DMAIC principles," *Journal of Evaluation in Clinical Practice*, vol. 23, no. 6, pp. 1401–1407, 2017.
- [30] G. Improta, G. Balato, C. Ricciardi et al., "Lean Six Sigma in healthcare: fast track surgery for patients undergoing prosthetic hip replacement surgery," *The TQM Journal*, vol. 31, no. 3, 2019.
- [31] A. A. Kuwaiti and A. V. Subbarayalu, "Reducing patients' falls rate in an academic medical center (AMC) using six sigma "DMAIC" approach," *International Journal of Health Care Quality Assurance*, vol. 30, no. 4, pp. 373–384, 2017.
- [32] A. A. Kuwaiti and A. V. Subbarayalu, "Reducing hospital-acquired infection rate using the six sigma DMAIC approach," *Saudi Journal of Medicine & Medical Sciences*, vol. 5, no. 3, pp. 260–266, 2017.
- [33] A. Al Kuwaiti, "Application OF six sigma methodology to reduce medication errors IN the outpatient pharmacy unit: a case study from the king fahd university hospital, Saudi Arabia," *International Journal for Quality Research*, vol. 10, no. 2, 2016.
- [34] G. Improta, G. Guizzi, C. Ricciardi et al., "Agile six sigma in healthcare: case study at santobono pediatric hospital," *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, 2020.
- [35] G. Improta, C. Ricciardi, A. Borrelli, A. D'alessandro, C. Verdoliva, and M. Cesarelli, "The application of six sigma to reduce the pre-operative length of hospital stay at the hospital Antonio Cardarelli," *International Journal of Lean Six Sigma*, vol. 11, no. 3, 2019.
- [36] J. Kalra and A. Kopargaonkar, "Quality improvement in clinical laboratories: a six sigma concept," *Pathology and Laboratory Medicine*, vol. 1, no. 1, pp. 11–20, 2016.
- [37] H. D. Koning, J. P. S. Verver, J. V. D. Heuvel, S. Bisgaard, and R. J. M. M. Does, "Lean six sigma in healthcare," *Journal for Healthcare Quality: Official Publication of the National Association for Healthcare Quality*, vol. 28, no. 2, pp. 4–11, 2006.
- [38] E. Montella, M. V. D. Cicco, A. Ferraro et al., "The application of Lean Six Sigma methodology to reduce the risk of healthcare-associated infections in surgery departments," *Journal of Evaluation in Clinical Practice*, vol. 23, no. 3, pp. 530–539, 2017.
- [39] G. Improta, G. Balato, M. Romano et al., "Lean Six Sigma: a new approach to the management of patients undergoing prosthetic hip replacement surgery," *Journal of Evaluation in Clinical Practice*, vol. 21, no. 4, pp. 662–672, 2015.
- [40] W. Scotton, R. Cobb, L. Pang et al., "Post-operative wound infection in salvage laryngectomy: does antibiotic prophylaxis have an impact?" *European Archives of Oto-Rhino-Laryngology: Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, vol. 269, no. 11, pp. 2415–2422, 2012.
- [41] A. M. Chen, L. M. Chen, A. Vaughan et al., "Tobacco smoking during radiation therapy for head-and-neck cancer is associated with unfavorable outcome," *International Journal of Radiation Oncology, Biology, Physics*, vol. 79, no. 2, pp. 414–419, 2011.
- [42] S.-C. Chuang, M. Jenab, J. E. Heck et al., "Diet and the risk of head and neck cancer: a pooled analysis in the INHANCE

- consortium,” *Cancer Causes & Control: Cancer Causes & Control*, vol. 23, no. 1, pp. 69–88, 2012.
- [43] T. F. S. Cunha, T. A. M. Soares, C. M. Z. F. Ribeiro, J. A. D. B. Almeida, S. S. M. Miguel, and D. A. E. D. B. C. André, “Risk factors for surgical site infection in cervico-facial oncological surgery,” *Journal of Cranio-Maxillofacial Surgery*, vol. 40, no. 5, pp. 443–448, 2012.
- [44] M. Hashibe, P. Brennan, S. Chuang et al., “Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the INHANCE consortium. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research,” *cosponsored by the American Society of Preventive Oncology*, vol. 18, no. 2, pp. 541–550, 2009.
- [45] A. R. Jethwa and S. S. Khariwala, “Tobacco-related carcinogenesis in head and neck cancer,” *Cancer and Metastasis Reviews*, vol. 36, no. 3, pp. 411–423, 2017.
- [46] N. Penel, C. Fournier, D. Lefebvre, and J. L. Lefebvre, “Multivariate analysis of risk factors for wound infection in head and neck squamous cell carcinoma surgery with opening of mucosa. Study of 260 surgical procedures,” *Oral Oncology*, vol. 41, no. 3, pp. 294–303, 2005.
- [47] L. Radoï and D. Luce, “A review of risk factors for oral cavity cancer: the importance of a standardized case definition,” *Community Dentistry and Oral Epidemiology*, vol. 41, no. 2, pp. e78–91, 2013.
- [48] R. Saulle, L. Semyonov, A. Mannocci et al., “Human papillomavirus and cancerous diseases of the head and neck: a systematic review and meta-analysis,” *Oral Diseases*, vol. 21, no. 4, pp. 417–431, 2015.
- [49] K. Loftus, T. Tilley, J. Hoffman, E. Bradburn, and E. Harvey, “Use of Six Sigma strategies to pull the line on central line-associated bloodstream infections in a neurotrauma intensive care unit,” *Journal of Trauma Nursing*, vol. 22, no. 2, pp. 78–86, 2015.
- [50] R. Simo and G. French, “The use of prophylactic antibiotics in head and neck oncological surgery,” *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 14, no. 2, pp. 55–61, 2006.
- [51] M. Garnier, C. Blayau, J.-P. Fulgencio et al., “[Rational approach of antibioprohylaxis: systematic review in ENT cancer surgery],” *Annales Françaises d’Anesthésie et de Réanimation*, vol. 32, no. 5, pp. 315–324, 2013.
- [52] Y. M. Haidar, P. B. Tripathi, T. Tjoa et al., “Antibiotic prophylaxis in clean-contaminated head and neck cases with microvascular free flap reconstruction: a systematic review and meta-analysis,” *Head & Neck*, vol. 40, no. 2, pp. 417–427, 2018.
- [53] J. T. Johnson, L. Y. Victor, E. N. Myers, R. R. Muder, P. B. Thearle, and W. F. Diven, “Efficacy of two third-generation cephalosporins in prophylaxis for head and neck surgery,” *Archives of Otolaryngology*, vol. 110, no. 4, pp. 224–227, 1984.
- [54] S. Egan, P. G. Murphy, J. P. Fennell et al., “Using Six Sigma to improve once daily gentamicin dosing and therapeutic drug monitoring performance,” *BMJ Quality & Safety*, vol. 21, no. 12, pp. 1042–1051, 2012.
- [55] C. Ricciardi, A. Sorrentino, G. Improta et al., “A health technology assessment between two pharmacological therapies through Six Sigma: the case study of bone cancer,” *The TQM Journal*, vol. 32, no. 6, 2020.
- [56] M. Pipan, “Quality improvement methodologies: PDCA cycle, RADAR matrix, DMAIC and DFSS,” 2010.
- [57] C. B. Mahesh, B. K. Ramakant, and V. S. Jagadeesh, “The prevalence of inducible and constitutive clindamycin resistance among the nasal isolates of staphylococci,” *Journal of Clinical and Diagnostic Research: Journal of Clinical and Diagnostic Research*, vol. 7, no. 8, pp. 1620–1622, 2013.
- [58] L. Lazzarini, M. Brunello, E. Padula, and F. D. Lalla, “Prophylaxis with cefazolin plus clindamycin in clean-contaminated maxillofacial surgery,” *Journal of Oral and Maxillofacial Surgery: Official Journal of the American Association of Oral and Maxillofacial Surgeons*, vol. 62, no. 5, pp. 567–570, 2004.
- [59] J. M. Heit, M. R. Stevens, and K. Jeffords, “Comparison of ceftriaxone with penicillin for antibiotic prophylaxis for compound mandible fractures,” *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology & Endodontics*, vol. 83, no. 4, pp. 423–426, 1997.
- [60] K. Dhanuthai, S. Rojanawatsirivej, W. Thosaporn et al., “Oral cancer: a multicenter study,” *Medicina Oral, Patología Oral y Cirugía Bucal*, vol. 23, no. 1, pp. e23–e29, 2018.
- [61] I. Petrovic, Z. U. Ahmed, A. Hay et al., “Sarcomas of the mandible,” *Journal of Surgical Oncology*, vol. 120, no. 2, pp. 109–116, 2019.
- [62] T. D. Shellenberger and E. M. Sturgis, “Sarcomas of the head and neck region,” *Current Oncology Reports*, vol. 11, no. 2, pp. 135–142, 2009.
- [63] A. Ketabchi, N. Kalavrezos, and L. Newman, “Sarcomas of the head and neck: a 10-year retrospective of 25 patients to evaluate treatment modalities, function and survival,” *British Journal of Oral and Maxillofacial Surgery*, vol. 49, no. 2, pp. 116–120, 2011.
- [64] S. Lester and W.-Y. Yang, “Principles and management of head and neck cancer,” *Surgery - Oxford International Edition*, vol. 33, no. 12, pp. 620–626, 2015.
- [65] L. Pecchia, N. Pallikarakis, R. Magiavere, and E. Iadanza, “Health technology assessment and biomedical engineering: global trends, gaps and opportunities,” *Medical Engineering & Physics*, vol. 72, pp. 19–26, 2019.
- [66] G. Improta, M. A. Russo, M. Triassi, G. Converso, T. Murino, and L. C. Santillo, “Use of the AHP methodology in system dynamics: modelling and simulation for health technology assessments to determine the correct prosthesis choice for hernia diseases,” *Mathematical Biosciences*, vol. 299, pp. 19–27, 2018.
- [67] E. Y. Bae, J. M. Hong, H. Y. Kwon et al., “Eight-year experience of using HTA in drug reimbursement: South Korea,” *Health Policy*, vol. 120, no. 6, pp. 612–620, 2016.
- [68] L. Maynou and J. Cairns, “What is driving HTA decision-making? Evidence from cancer drug reimbursement decisions from 6 European countries,” *Health Policy*, vol. 123, no. 2, pp. 130–139, 2019.
- [69] M. Shayne, E. Culakova, M. S. Poniewierski et al., “Risk factors for in-hospital mortality and prolonged length of stay in older patients with solid tumor malignancies,” *Journal of Geriatric Oncology*, vol. 4, no. 4, pp. 310–318, 2013.
- [70] R. Grossman, D. Mukherjee, D. C. Chang et al., “Preoperative charlson comorbidity score predicts postoperative outcomes among older intracranial meningioma patients,” *World Neurosurgery*, vol. 75, no. 2, pp. 279–285, 2011.

Research Article

A Machine Learning-Based Prediction Model for Preterm Birth in Rural India

Rakesh Raja , **Indrajit Mukherjee**, and **Bikash Kanti Sarkar**

Department of Computer Science & Engineering, Birla Institute of Technology, Mesra, Ranchi, India

Correspondence should be addressed to Rakesh Raja; rajarakeshchauhan@gmail.com

Received 20 November 2020; Revised 16 March 2021; Accepted 17 March 2021; Published 15 June 2021

Academic Editor: Giovanni Improta

Copyright © 2021 Rakesh Raja et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preterm birth (PTB) in a pregnant woman is the most serious issue in the field of Gynaecology and Obstetrics, especially in rural India. In recent years, various clinical prediction models for PTB have been developed to improve the accuracy of learning models. However, to the best of the authors' knowledge, most of them suffer from selecting the most accurate features from the medical dataset in linear time. The present paper attempts to design a machine learning model named as risk prediction conceptual model (RPCM) for the prediction of PTB. In this paper, a feature selection approach is proposed based on the notion of entropy. The novel approach is used to find the best maternal features (responsible for PTB) from the obstetrical dataset and aims to predict the classifier's accuracy at the highest level. The paper first deals with the review of PTB cases (which is neglected in many developing countries including India). Next, we collect obstetrical data from the Community Health Centre of rural areas (Kamdara, Jharkhand). The suggested approach is then applied on collected data to identify the excellent maternal features (text-based symptoms) present in pregnant women in order to classify all birth cases into term birth and PTB. The machine learning part of the model is implemented using three different classifiers, namely, decision tree (DT), logistic regression (LR), and support vector machine (SVM) for PTB prediction. The performance of the classifiers is measured in terms of accuracy, specificity, and sensitivity. Finally, the SVM classifier generates an accuracy of **90.9%**, which is higher than other learning classifiers used in this study.

1. Introduction

Preterm birth (PTB) is a serious public health problem that adversely affects both families and the society [1]. It is a leading cause of neonatal mortality and morbidity across the world and also the second major cause of child deaths under the age of five years [2]. Over the past two decades, PTB has been a significant research study in healthcare domain. Pregnancy and childbirth unlocked the door for medical experts and researchers to explore various effective strategies to reduce preterm birth in women having pregnancy-related complications. These strategies include healthcare services given to all pregnant women to control PTB and any medical interventions aimed to enhance the knowledge of women on early indications of pregnancy complications [3, 4]. The maternal history of a pregnant woman is a key part of the neonatal studies for providing certain clinical treatments to newborn babies regarding their health, disease, care, and outcomes. Newborn babies are very special. They do not

have any previous medical background, and their early neonatal path is directly connected to the maternal history of their mothers [5–7]. The healthcare services also incorporate the arrangements of essential social and economic support for women before, during, and after pregnancy including educational, medical, and other training programs that facilitate healthy motherhood.

In general, treatments of diseases (including PTB) are made by the physicians based upon their knowledge (experience). However, on the one hand, manual diagnosis may not be often right as physician's experience varies from expert to expert. On the other hand, manual treatment is a time-consuming job. Further, shortage of medical experts is increasing everyday with population explosion and developing countries like in India, large number of women belong to lower or middle income families. They do not get proper healthcare facilities or awareness regarding health education to know about any complication that arises during pregnancy, especially in rural area. Further, people often are

afraid of doctors' prescription since doctors in most cases misguide the patients suggesting unnecessary tests (like double marker test, fetal echocardiography, urine test, and FT4 test which are used to determine any pregnancy complications) which are very expensive. Also, doctor's appointment fees are mostly on higher side. Besides, doctors could sometime diagnose the cases wrongly. After all, preterm delivery is the most critical issue in Gynaecology and Obstetrics and a major health concern for every pregnant woman. It may require several ultrasound sonography (USG) tests in addition to doctor's appointment fee for diagnosing high-risk patients, and these altogether may amount huge expense that may be beyond the income limit of many families. So, designing the computerized system (i.e., e-healthcare system) for birth prediction from past diagnosis data is the essential solution for quick and accurate decision to be taken for any adverse pregnancy outcome in order to save lives and cost.

Notably, a pioneering renovation is taking place in the Obstetrical community due to the advancement in technology and digitization of medical records. Data analytics is one of the most promising tool for research and development in the area of medicine [8–15]. Nowadays, machine learning techniques (e.g., neural networks, support vector machine, logistic regression, and Decision Trees) are playing important role in designing the disease predictive model to address the growing needs of human experts in the medical world [16–20]. However, medical datasets are highly imbalanced, conflicting in nature, and uncertain. So, designing the effective intelligent model for medical datasets is a challenging task. PTB dataset is one such clinical dataset. Numerous predictive models based on standard intelligent methods have been introduced by the researchers for prediction of PTB [21]. However, they usually suffer from several drawbacks like lack of understandability and inefficiency in making quick and correct decision. Further, early detection and diagnosis play important role in controlling such complications. Symptoms (text) based machine intelligent models may play vital role in early detection of such cases. The delay in receiving the clinical judgement for preterm delivery increases the risk of pregnancy complications which in turn increases the risk of prenatal mortality. Due to its direct association with prenatal mortality, neonatal health is also very important in the obstetrical community [7]. According to the UNICEF study released in 2015, 35% of neonatal death is due to PTB. The rate of PTB in rural areas of most developing countries is increasing due to lack of health facilities and insufficient number of healthcare workers.

In light of these considerations, the present study aims to design a novel conceptual model (by employing machine learning techniques) and its implementation for detection of PTB in pregnant women. In fact, the system can be used as a decision support system to assist the medical staff and healthcare workers for predicting premature delivery. More specifically, the present study focuses on novel feature selection (entropy-notion) approach to identify the most important maternal features (text-based symptoms)

responsible for preterm delivery and aims to predict the classification accuracy.

The remaining sections of the paper are organized as follows. Section 2 describes the basic concept of PTB and feature selection. Section 3 elaborates the related work that has been carried out to predict PTB. Section 4 describes the methodology of this research. The experimental design and results are presented in Section 5. Finally, Section 6 deals with conclusion and future scopes.

2. Background of the Present Research

2.1. Preterm Birth (PTB): A Comprehensive Overview. Preterm or premature birth is defined as birth, for any reason, occurring before 37 completed weeks (or less than 259 days) of pregnancy. Every year, about fifteen million babies are born prematurely (before 37 completed weeks of gestation), and this is nearly equal to one-tenth of all babies around the world [22]. According to the WHO reports studied in 2005, 12.9 million births or 9.6% of all births across the world occurred prematurely [23]. The rate of preterm birth, however, significantly varies across the world. Preterm birth reflects the most prominent reason for neonatal morbidity and mortality [24].

2.1.1. Categorization of PTB. PTB can be classified into different categories based on gestational age at birth. The gestational age is defined as the time from the first day of the last menstrual period (LMP) of a woman to birth [21]. The four categories of PTB are as follows:

- (i) *Extreme PTB (under 28 Weeks).* It is the birth that takes place before 28 weeks of pregnancy
- (ii) *Very PTB (28 to 32 Weeks).* It is the birth that takes place between 28 and 32 weeks of pregnancy
- (iii) *Moderate PTB (32 to 34 Weeks).* It is the birth that takes place between 32 and 34 weeks of pregnancy
- (iv) *Late PTB (34 to 37 Weeks).* It is the birth that takes place between 34 and 37 weeks of pregnancy

2.1.2. Medical Terminologies. For the purpose of clarity of the present study, the used terminologies are illustrated in Table 1.

2.1.3. Health Impact of PTB. PTB is the main risk factor for newborn mortality and morbidity. It is a leading cause of neonatal mortality and morbidity across the world and also the second major cause of child deaths under the age of five years [25]. It arises between 5 and 10% of all deliveries and involves 70% of neonatal mortality and up to 75% of neonatal morbidity [26]. Premature infants are more likely to suffer than normal birth and are at higher risk of brain paralysis, sensory impairment, respiratory failure, and so on. More than \$13 billion of premature cost for maternity service is anticipated only in the USA [27, 28]. Most survivors of PTB face serious problems, often a lifetime of disability, including learning disabilities, visual, and hearing problems.

TABLE 1: Definitions used in the present study.

Terminology	Description
Antenatal care	Antenatal care (ANC) refers to the fundamental, clinical, and nursing care suggested for ladies during pregnancy
Neonate	A neonate or a newborn infant is a child under 28 days of age
Neonatal death	A death during the first 28 days of life (0–27 days) is termed as a neonatal death
Live birth	A birth at which a child is born alive is termed as live birth
Term birth	A birth at the end of a normal duration of pregnancy between 37 and 40 weeks of gestation is termed as term birth
Maternal death	A maternal death is the death of a woman while pregnant or within 42 days of termination of pregnancy
Stillbirth	Stillbirth is the delivery, after the 20th week of pregnancy, of a baby who has died
Abortion	Termination of a pregnancy either medically or induced
Miscarriage	Natural loss of pregnancy during first trimester
Gestational age	Gestational age (GA) refers to the time from the first day of a woman's last menstrual period to birth

In fact, babies born premature have more health problems compared with babies born at term birth. Term birth refers to babies that are born at 37 to 40 weeks of gestation. Furthermore, babies born at preterm are reported to be at an elevated risk of long-term health problems [29]. Unfortunately, after many years of research in obstetrics, yet the rate of PTB has not decreased [30]. Birth weight is generally associated with PTB and results in its own categorization. Usually, birth weight is simpler to measure precisely and is a first estimation of gestational age. Obviously, the most challenging issue in Gynaecology and Obstetrics is how to control the preterm delivery in pregnant women.

2.2. Feature Selection (FS). The term feature selection in the machine learning, also known as feature subset selection, refers to the process of selecting a subset of excellent features during construction of the predictive model. The presence of redundant and irrelevant features in any datasets (especially in medical datasets) can reduce the accuracy of the model's prediction and also have the negative impact on the performance of the model. The main goal of any feature selection method is to select the best subset of features by removing redundant and irrelevant features from the datasets in order to reduce the training time and enhance the classifier's predictive performance. In fact, feature selection is typically used as a preprocessing step in data mining. There are three standard approaches of the feature selection algorithm, namely, filter method, wrapper method, and embedded method. For more details about feature selection, one may refer to [31–33].

- (i) *Filter Method.* The filter method measures the relevance of features based on the nature of data. The selection of features is independent of the classifiers used. The filter method is much faster compared with the wrapper method and provides an average accuracy for all the classifiers used. Some of the examples of filter methods are information gain, chi-square test, variance threshold, and so on.
- (ii) *Wrapper Method.* The wrapper method finds the best subset of features based on a specific machine learning algorithm that we are trying to fit on a given dataset. The evaluation criteria are simply the

predictive power of the particular classifier. The wrapper method has higher performance accuracy compared with the filter method but requires more computational time to find best features for a dataset with high-dimensional features. Some of the examples of wrapper methods are forward selection, backward elimination, genetic algorithms, and so on.

- (iii) *Embedded Method.* The embedded method incorporates the advantages of both filter and wrapper methods. In this approach, feature selection is done during the process of model training and is usually unique to particular learning classifiers. This approach basically determines the importance of feature, i.e., which features to accept and which to reject, while making a prediction. The most typical embedded technique is the decision tree algorithm. This method typically falls somewhere between the filter method and wrapper method in terms of time complexity. Some of the examples of embedded methods are lasso regression, ridge regression, elastic net, and so on.

3. Related Works

This section focuses mainly on the existing methodologies related to prediction of PTB using machine learning, statistical analysis, and data mining techniques. Some of them are discussed in this section. The study of Mercer et al. [34] was designed to develop a risk-score-based model for predicting PTB. The model can be trained using a multivariate logistic regression technique to explore various risk factors using clinical data available between 23 and 24 weeks' gestation. Goodwin et al. employed the machine learning model to generate 520 predictive rules for PTB with the application of data mining techniques [35]. The study in [36] discussed the deep learning models for predicting preterm delivery using existing electronic medical records (EMRs) of mothers available in healthcare centres.

Weber et al. [37] performed a cohort study to predict spontaneous preterm. The prediction of PTB was performed using numerous classifiers, namely, K-nearest neighbours, lasso regression, and random forests. This study has taken

into the consideration of demographic, race-ethnicity, and maternal characteristics. Mailath-Pokorny et al. [38] explored the predictive features for preterm delivery that occurs within 2 days after admission and before 224 days of gestation using the multivariate logistic regression model. The predictive features considered are age of the mother, gestational age during admission, maternal history, vaginal bleeding, cervical length, preterm history, and preterm premature rupture of membranes (PPROM) in their study. Son and Miller presented a prediction model for PTB using cervical length measurement in women with a singleton gestation. To accomplish better predictive performance, they attempted to determine the best cut points of cervical length [39].

Elaveyini et al. [40] explored the major risk factors of preterm birth using artificial neural networks. PTB prediction was based on the feed-forward backpropagation algorithm. Over the past decades, majority of research studies have been done to enhance the accuracy of prediction of PTB [41]. Researchers are continually making their best efforts to analyse and explore the principal risk factors for preterm delivery [42–44]. The present article focuses on the machine learning approaches for prediction of birth cases in rural community.

3.1. Shortcomings in the Existing Clinical Models. In recent years, using feature selection approach, a significant number of clinical prediction model have been developed to improve the accuracy of learning models. However, to the best of authors' knowledge, most of them suffer from selecting the most accurate features from the medical dataset in linear time. Hence, there is a scope for improving the performance of machine learning classifiers and reducing learning time.

3.2. Novel Contribution. A novel feature selection approach based on the notion of entropy is introduced in this study to address the identified issues of the existing models. The key role of the novel approach is to find the subset of optimal features from the medical dataset in order to improve the prediction's accuracy and ultimately reduce the machine learning time.

4. Research Methodology

4.1. Objective. The finding of this research study can be utilized to fulfill the three following main objectives:

- (i) A machine learning-based risk prediction conceptual model (RPCM) for PTB can be introduced with the help of novel feature selection approach using entropy-notion to predict the birth cases (TB and PTB) from the obstetrical records.
- (ii) The suggested approach is used to identify the excellent (text-based symptoms) features responsible for PTB. Furthermore, medical experts'

(physicians and obstetricians) opinions are also considered through review of medical records of patients and survey analysis. The model can be extended to select the regions for pregnancy consultation.

- (iii) The predictive model can be beneficial for rural India to identify the important maternal features in order to predict the possibility of PTB in the gestation of women. This information can support rural medical staff for taking effective decisions for adverse pregnancy outcome—that aim to reduce the diagnosis cost.

4.2. The Proposed Feature Selection Approach Based on the Notion of Entropy. According to the study in [45], attributes having strong correlation cannot be the part of feature subset. Besides, more the attributes are independent among themselves and more information gain they will have which would eventually give better outcomes over unseen data. The present research focuses on medical (obstetrical) datasets which are more sensitive in nature, so feature selection approach is more effective for such datasets. In light of this point, a feature selection (entropy-notion) approach is presented here to extract the most relevant features from obstetrical (term-preterm) dataset. These features are utilized to classify all birth cases into term birth and PTB. A conceptual model of the proposed approach is shown in Figure 1.

The proposed approach is stated as follows:

- (i) Suppose that D is a medical dataset having n attributes, say A_i for $i = 1, 2, 3, \dots, n$.

Let F_0 denote a set of features in the original dataset D .

Initially, $F_0 = \{A_1, A_2, A_3, \dots, A_n\}$.

Since D is divided into three distinct subsets as D_1 , D_2 , and D_3 , so after applying the proposed approach, we get three feature subsets, namely, F_1 , F_2 , and F_3 from these data subsets.

F is considered as a resultant feature set derived from F_1 , F_2 , and F_3 . Initially, $F_k = F_0$ for $k = 1, 2, 3$.

Let P be a classification problem described by a set of n attributes, say A_i for $i = 1, 2, 3, \dots, n$ and also consider that F represents the set of features derived from the original dataset.

Initialize, $F = F_0 = \{A_1, A_2, A_3, \dots, A_n\}$.

for each data subset $D_i \in D$; where $i = 1, 2, 3$

do

for each attribute $A_i \in F_0$

do

Calculate Gain (S, A_i)/information gain for A_i

Using formula stated below,

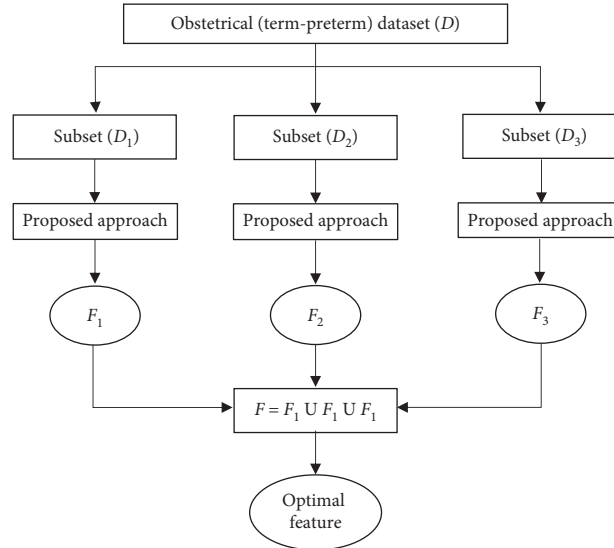


FIGURE 1: Conceptual model for feature selection approach.

Gain $(S, A_i) = \text{Entropy}(S) - \sum_{v_j \in A_i} (|S_{v_j}|/|S|) \text{Entropy}(S_{v_j})$, where v_j denotes values of attribute A_i and $\text{Entropy}(S) = -\sum p_m \log_2 p_m$, where S represents the number of instances in P and p_m is the nonzero probability of s_m instances (out of S) belonging to class m , out of c classes.

end for

Compute

$r = (\text{Max_Gain}(S, A_i) - \text{Min_Gain}(S, A_i))/n$, where $i = 1, 2, 3, \dots, n$.

//Here, r is considered as a threshold value for selecting features

for each attribute $A_i \in F_0$

do

if $\text{Gain}(S, A_i) < r$

then

update $F_k = F_k - \{A_i\}$ //removing A_i from F_k

end if

end for

end for

$F = F_1 \cup F_2 \cup F_3$ //including all attributes of F_1, F_2 , and F_3 .

Note. The proposed feature selection approach in this study is a form of the filter method and is implemented in Java-1.4.

Time Complexity. The algorithm is simple and easy to understand. The running time of an algorithm is $O(n)$, where n is the number of attributes in the dataset.

4.3. The Proposed Framework: Risk Prediction Conceptual Model (RPCM). Based on novel feature selection (entropy-notation) approach and several studies in [46–49], RPCM is carefully designed to predict the risk of PTB in pregnant

women. The workflow of the framework consisting of three stages (Stage-I, Stage-II, and Stage-III) is depicted in Figure 2, and then its each component is detailed.

4.3.1. Key Components of the Proposed Model. The proposed model consists of some key components, namely, healthcare centre, maternal and neonatal records, data preprocessing, machine learning, and birth outcome. Each of these is discussed as follows:

- (i) **Healthcare Centre.** A healthcare centre is a part of a network of hospitals employed by a group of general physicians, nurses, and healthcare professionals that provide healthcare facilities to people in a certain area. In addition to standard medical treatments, one of the main goals of the primary healthcare centre is maternal care during pregnancy especially in rural India. This is because people from rural India avoid contacting healthcare professionals and practitioners for pregnancy care which increases the cases of maternal and neonatal deaths.
- (ii) **Patient Survey.** A comprehensive care to mother and child is primarily concerned to all healthcare systems in India. The term survey describes any study that consists of requesting people to respond queries. This entails researcher-developed questionnaires and personal interviews with pregnant women during their antenatal care visits.
- (iii) **Maternal and Neonatal Records.** Maternal and neonatal records play a vital role in deciding the way healthcare services are provided, accessed, and affected by health outcomes. It stores the statistical reports describing the use of prenatal services, maternal risk factors, and birth outcomes for all patients residing in rural area. PTB is one of the most frequent complication of pregnancy. It occurs due to several medical reasons and is affected by

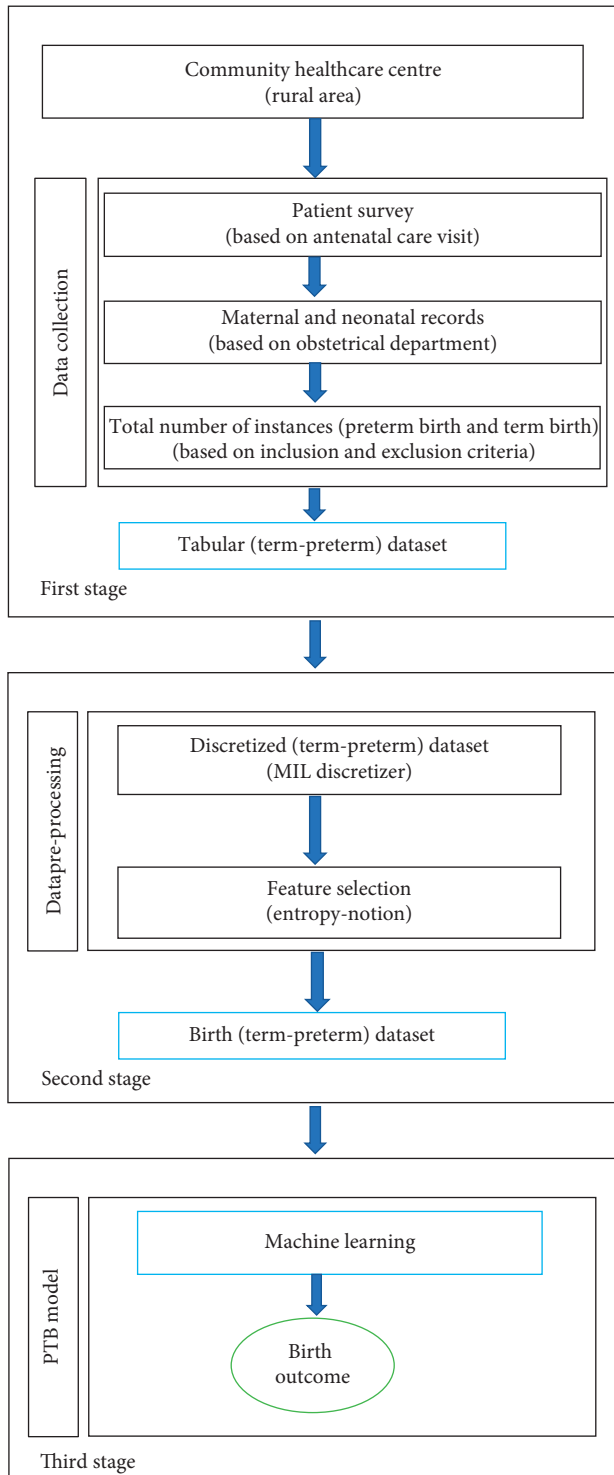


FIGURE 2: Framework of the proposed model.

some of the important maternal features based on human experts (experience) and several research studies [50–53]. These maternal features are critical in nature to predict cases of PTB. The total number of birth instances is taken from the obstetrical data.

- (iv) *Data Discretization.* A technique of converting continuous values of attribute into a finite set of

intervals and associating a new discrete value with each interval is known as data discretization. Since any classifiers prefer to handle discrete values rather than continuous values for the learning process, data discretization plays a crucial role in the process of machine learning. The study in [54] suggests that data discretization improves the quality of discovered knowledge, and it is based on the concept of information theory.

- (v) *Feature Selection.* One of the core concepts in machine learning is the feature selection. Feature selection is the process of selecting those features from the input datasets which highly impact the performance of the predictive model. The present study focuses on feature selection approach based on entropy notion as already discussed in Section 4.2.
- (vi) *Data Preprocessing.* The tabular dataset collected from obstetrical data is preprocessed and converted into a normalized form with the help of MIL discretizer [55, 56].
- (vii) *Machine Learning (ML).* The present study focuses on applying machine learning algorithms [46, 49] for PTB prediction. ML is a method of data analysis that automates analytical model building. Classification is one of the most popular approaches for applying ML methods (e.g., DT, LR, and SVM). These techniques are used in medical domain for classification, prediction, and diagnosis purposes.
- (viii) *Birth Outcome.* This component is very crucial in preventing preterm delivery in pregnant women during antenatal care clinics. The predicted birth outcome can also be used to properly analyse the key maternal features responsible for PTB.

4.4. Details of Stage-I. The main role of the first stage of framework is to collect obstetrical data from the Community Healthcare Centre, and it is detailed in this section.

4.4.1. Study Design. The study was conducted in the Community Health Centre, Kamdara (Gumla), situated in rural area of Jharkhand, during a period from July 2018 to September 2020. The hospital provides obstetric and gynaecological services to all categories of women, whether registered for antenatal care or referred. The approval for the study was taken from the Institutional Ethics Committee.

Selection Criteria. The selection of patients (women) depends on the following inclusion-exclusion criteria:

Inclusion criteria include the following:

- (i) Women registered for ANC and having birth at the Community Health Centre
- (ii) Women having birth occurring at the gestational age of 28 weeks or more
- (iii) Women who delivered a live birth

Exclusion criteria include the following:

- (i) Women having still birth
- (ii) Women having birth of twins
- (iii) Women referred to other hospitals

4.4.2. Data Collection. The basic step of Stage-I is to collect data based on patient survey and maternal records available in the obstetrics department. Initially, 1800 records were collected during a research period. Then, 1300 records were selected for further study based upon inclusion-exclusion criteria. The collected records include all instances of term birth and PTB. A manual analysis is performed to select all maternal features which are involved during pregnancy (based on medical experts' opinion and several research studies) [51, 52, 57, 58]. The description of the obstetrical dataset (original) after data collection is summarized in Table 2.

Initially, all instances are in a raw-form which are compiled into a tabular-form using MS Excel program. As a result, a tabular (term-preterm) dataset is prepared for the research purpose. The tabular (term-preterm) dataset used in this work is a binary class dataset.

The feature values in this dataset are of the form-string, integer, and continuous. The tabular (term-preterm) dataset consists of 1300 instances, composed of thirty-six different features which are taken into consideration before, during, and after pregnancy. These features are listed in Table 3. The questionnaire used for data entry during patient survey was mainly focused on their background details, medical history, previous pregnancy details, current pregnancy details, baby details, and medical disorders in current pregnancy.

4.5. Description of Stage-II. The collected data from tabular (term-preterm) dataset are preprocessed at the second stage of the framework. This stage deals with two main operations, namely, data discretization and feature selection.

4.5.1. Data Discretization. During data preprocessing, tabular (term-preterm) dataset is converted into a normalized form with the help of data discretization process. This gives a discretized (term-preterm) dataset. This dataset is utilized to select most accurate features by applying suggested feature selection approach based on the notion of entropy. The initial statistics of discretized (term-preterm) dataset is shown in Table 4.

In reality, attributes of any medical dataset may contain mixture of string, continuous, outliers, and missing data. Many classifiers cannot handle continuous attributes but each of them can operate on discretized attributes [55]. Besides, performance of classifiers can be significantly improved by replacing continuous attributes with its discretized values. Depending upon the amount of missing data and the criticality of the feature in which the data is missing, it may impact the accuracy of prediction. In this study, the missing value in any feature is replaced with the mean value of that feature, and minimum information loss (MIL) data discretizer [12, 54, 59] is employed here for data processing,

which make data compatible with the machine learning algorithm.

4.5.2. Feature Selection. After that, the proposed feature selection approach is taken into consideration to select the most probable features (responsible for PTB) from the discretized (term-preterm) dataset. As a result, seventeen different features are selected from this dataset. These maternal features (listed in Table 5) are also considered as major risk factors for PTB as suggested by medical experts and several research studies. Then, a final birth (term-preterm) dataset, consisting of these selected features, is prepared for the last stage of framework. The birth dataset also contains 1300 instances of term birth and PTB.

4.6. Description of Stage-III. Finally, a machine learning-based prediction model for PTB is built at this stage. This section describes the actual construction of the suggested system.

4.6.1. Machine Learning PTB Model. The aim of this research is to find a suitable classifier which can predict the PTB with more accuracy. The three classifiers, namely, decision tree (DT), logistic regression (LR), and support vector machine (SVM) are used in this analysis. The method of selecting classifier in this study is illustrated in Figure 3. Model fitting was carried out by dividing the input dataset into training dataset and test dataset at a ratio of 70% and 30%, respectively. The training set is used in learning phase and test set is used in prediction phase, to determine the best model. Researchers may find ample information about several machine learning classifiers from the articles [60–63].

4.6.2. Evaluation of Machine Learning Classifiers. The empirical measures can be extracted from the confusion matrix in order to evaluate the performance of the learning classifier [64]. A confusion matrix shows the accuracy of the solution to a classification problem. Table 6 depicts the confusion matrix, which summarizes the number of instances predicted correctly or incorrectly by a classification model.

Furthermore, the other parameters used to measure the classifier's performance are correct classification rate (CCR) or accuracy, true positive rate (TPR) or sensitivity, true negative rate (TNR) or specificity, false positive rate (FPR), false negative rate (FNR), precision, recall, and F1 score. A formal definition of these performance metrics is shown in Table 7.

5. Experimental Design and Results

5.1. Experimental Design. A birth (term-preterm) dataset with 1300 patients' observations is obtained in order to perform the experiment. The experiment is carried out with the help of Python and Scikit-Learn library or under WEKA toolbox (<http://www.cs.waikato.ac.nz/ml/weka>). The observations in the birth dataset are carefully reviewed for prediction of birth cases. This is in fact a binary class dataset in

TABLE 2: Summary of the obstetrical (term-preterm) dataset.

Problem name	Number of features	Number of classes	Number of instances
Birth case	36	2	1300

TABLE 3: Maternal features associated with PTB.

S. no.	Feature ID	Feature name
1	PID	Patient identification
2	WA	Woman age
3	LMP	Last menstrual period
4	EDD	Estimated delivery date
5	G	Gravida
6	P	Parity
7	A	Abortion
8	L	Living
9	EL	Educational level
10	H	Height
11	W	Weight
12	BMI	Body mass index
13	BP	Blood pressure
14	HB	Hemoglobin
15	ANC	Antenatal care visit
16	ADD	Actual delivery date
17	OH	Obstetric history
18	PCS	Previous caesarean section
19	GA	Gestational age
20	BW	Birth weight
21	GDM	Gestational diabetes mellitus
22	FHR	Fetal heart rate
23	MG	Multiple gestation
24	ND	Normal delivery
25	MH	Previous medical history
26	LBW	Low birth weight
27	ASPX	Asphyxia
28	HT	Hypertension
29	PE	Preeclampsia
30	LV	Live birth
31	SB	Still birth
32	OB	Obesity
33	AN	Anemia
34	TH	Thyroid
35	NS	Neonatal status
36	PTB	Preterm birth

TABLE 4: Summary of discretized (term-preterm) dataset.

Outcome	<i>N</i>
Number of features	36
Number of classes	2
Total instances	1300
Term birth	991
Preterm birth	309

which all births occurring between 28th to 37th weeks are termed as PTB class with label “1” whereas all births after 37th weeks are termed as term birth (TB) class with label “0.” According to the study, around 24% of the findings in the

TABLE 5: List of excellent features in discretized (term-preterm) dataset.

Feature code	Feature name	Feature type
WA	Woman age	Numeric
PT	Parity	Numeric
GD	Gravida	Numeric
BMI	Body mass index	Ordinal
ANC	Antenatal care visit	Numeric
GA	Gestational age	Numeric
FHR	Fetal heart rate	Numeric
BP	Blood pressure	Ordinal
HB	Hemoglobin	Numeric
GDM	Gestational diabetes mellitus	Binary
PE	Preeclampsia	Binary
HT	Hypertension	Binary
OH	Obstetric history	Binary
EL	Education level	Ordinal
CS	Previous caesarean section	Binary
MH	Previous medical history	Binary
PTB	Preterm birth (target variable)	Binary

dataset are of PTB with label “1” and remaining 76% are of TB with label “0.” Hence, PTB class is dominated by TB class, and we can say that PTB is the minority class and TB is the majority class. Therefore, there is a need of a good sampling technique for medical datasets [24, 52]. In this context, synthetic minority oversampling technique (SMOTE) is used to balance the target dataset [65]. This can be achieved by replicating the PTB cases until it reaches approximately 50% of the dataset. This gives a new balanced (term-preterm) dataset.

5.2. Results and Discussion. A total of 1300 patients (women) were selected in this study based on inclusion-exclusion criteria. Out of 1300 pregnant women, 309 women were having preterm birth and rest 991 women were having term birth. Thus, the incidence of PTB is 23.78% of total pregnant women. In this work, the performance of DT, LR, and SVM classifiers is evaluated in terms of accuracy, specificity, and sensitivity [66]. With these indicators, it is possible to compare the proposed model performance with three classifiers. Tables 8 and 9 present the performance metrics of classifiers for the original dataset and balanced dataset, respectively.

Based on the results shown in Tables 8 and 9, we can observe that the accuracy of three different classifiers is roughly around 85%. With respect to the original dataset, the accuracy of SVM is 86.1% which is highest, followed by LR and DT. The results were additionally improved (after applying SMOTE) with the balanced dataset. The accuracy of SVM classifier in the balance dataset increases from 86.1% to

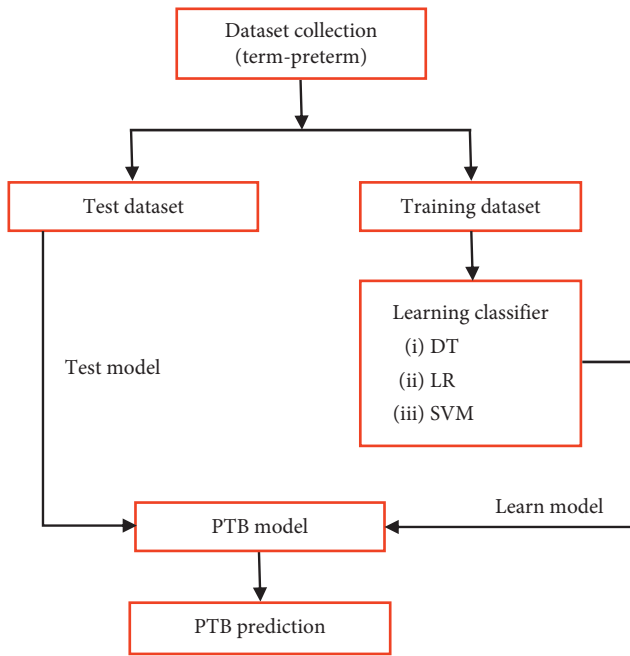


FIGURE 3: A conceptual PTB prediction model.

TABLE 6: Confusion matrix.

	Predictive positive	Predictive negative
Actual positive	True Positive (TP)	False Negative (FN)
Actual negative	False Positive (FP)	True Negative (TN)

TABLE 7: Performance metrics for machine learning classifiers.

Metrics	Formula
CCR	$((TP + TN)/(TP + FP + FN + TN))\%$
TPR	$TP/(TP + FN)$
TNR	$TN/(TN + FP)$
FPR	$FP/(TN + FP)$
FNR	$FN/(TP + FN)$
Precision	$TP/(TP + FP)$
Recall	$TP/(TP + FN)$
F_1 score	$2 * TP/(2 * TP + Fp + FN)$

TABLE 8: Performance metrics of the classifiers—original dataset.

Classifiers	Accuracy	Sensitivity	Specificity
DT	0.777	0.702	0.930
LR	0.841	0.863	0.971
SVM	0.861	0.801	0.702

TABLE 9: Performance metrics of the classifiers—balanced dataset.

Classifiers	Accuracy	Sensitivity	Specificity
DT	0.796	0.713	0.972
LR	0.872	0.832	0.954
SVM	0.909	0.891	0.783

90.9% compared with original dataset. In summary, the SVM model is the best classifier in the experiment.

6. Conclusion and Future Scope

In this study, the proposed model (RPCM) can be used for prediction of PTB based on excellent features (text-based symptoms) available in obstetrical data. The work focuses on feature selection (entropy-notion) approach by applying machine learning classifiers (DT, LR, and SVM) in order to classify all birth cases into term birth and PTB. Comparing the performances of the classifiers, it is evident that SVM classifier is the most suitable classifier as it achieves an accuracy of 90.9%. According to the findings of this study, the identified risk factors (excellent features) will be helpful in the prediction of PTB, especially in rural community. The developed model supports the decision-making process in maternity care by identifying and alerting the pregnant women at risk of preterm delivery thereby preventing possible complications, reducing the diagnosis cost, and ultimately minimizing the risk of PTB. The present system can be regarded as a successful innovation in Obstetrics to give clinical support to patients during pregnancy consultations. In particular, RPCM claims to assist healthcare professionals to make effective and timely decisions without consulting specialists directly.

The limitation of the present research is that the risk factors for PTB are limited in size and dataset is small, which could be increased to improve the performance of the PTB prediction in the future studies. However, expert knowledge and clinical judgement may still be needed to interpret this risk and take appropriate action in individual cases.

Data Availability

The data used to support the finding of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy and ethical restrictions of Institutional Ethics Committee.

Disclosure

The content of this paper represents the views of the authors and do not necessarily reflect the views of the Community Health Centre.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the staff of Department of Gynaecology and Obstetrics at Community Health Centre (Kamdara, Jharkhand) for their efforts in the completion of this research work. The authors gratefully thank Dr. Ruchi Bhushan, MBBS, MS (OBG), for her valuable comments and helpful discussions. The authors would also like to convey their sincere gratitude to Mrs. Priyanka and Ms. Prerna for their neverending support and motivation.

References

- [1] H. Blencowe, S. Cousens, D. Chou et al., "Born Too Soon: the global epidemiology of 15 million preterm births," *Reproductive Health*, vol. 10, no. 1, p. S2, 2013.
- [2] L. Liu, S. Oza, D. Hogan et al., "Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis," *The Lancet*, vol. 385, no. 9966, pp. 430-440, 2015.
- [3] B. Govindaswami, P. Jegatheesan, M. Nudelman, and S. R. Narasimhan, "Prevention of prematurity," *Clinics in Perinatology*, vol. 45, no. 3, pp. 579-595, 2018.
- [4] L. J. E. Meertens, P. van Montfort, H. C. J. Scheepers et al., "Prediction models for the risk of spontaneous preterm birth based on maternal characteristics: a systematic review and independent external validation," *Acta obstetrica et gynecologica Scandinavica*, vol. 97, no. 8, pp. 907-920, 2018.
- [5] M. J. Perez, J. J. Chang, L. A. Temming et al., "Driving factors of preterm birth risk in adolescents," *American Journal of Perinatology Reports*, vol. 10, no. 3, pp. e247-e252, 2020.
- [6] S. Shrestha, S. S. Dangol, M. Shrestha, and R. P. B. Shrestha, "Outcome of preterm babies and associated risk factors in a hospital," *Journal of the Nepal Medical Association*, vol. 50, no. 180, 2010.
- [7] C. Catley, M. Frize, R. C. Walker, and D. C. Petriu, "Predicting preterm birth using artificial neural networks," in *In Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*, pp. 103-108, Dublin, Ireland, June 2005.
- [8] A. Belle, R. Thiagarajan, S. M. R. Sorousmehr, F. Navidi, D. A. Beard, and K. Najarian, "Big data analytics in healthcare," *BioMed Research International*, vol. 2015, Article ID 370194, 16 pages, 2015.
- [9] R. Raja, I. Mukherjee, and B. K. Sarkar, "A systematic review of healthcare big data," *Scientific Programming*, vol. 2020, Article ID 5471849, 15 pages, 2020.
- [10] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *IEEE Access*, vol. 5, pp. 8869-8879, 2017.
- [11] A. Mansoul, B. Atmani, and S. Benbelkacem, "A hybrid decision support system: application on healthcare," 2013, <http://arxiv.org/abs/1311.4086>.
- [12] B. K. Sarkar, "A two-step knowledge extraction framework for improving disease diagnosis," *The Computer Journal*, vol. 63, no. 3, pp. 364-382, 2020.
- [13] M. Seera and C. P. Lim, "A hybrid intelligent system for medical data classification," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2239-2249, 2014.
- [14] A. Bhardwaj and A. Tiwari, "Breast cancer diagnosis using genetically optimized neural network model," *Expert Systems with Applications*, vol. 42, no. 10, pp. 4611-4620, 2015.
- [15] P. J. Lisboa and A. F. G. Taktak, "The use of artificial neural networks in decision support in cancer: a systematic review," *Neural Networks*, vol. 19, no. 4, pp. 408-415, 2006.
- [16] T. M. Mitchell, "Machine learning and data mining," *Communications of the ACM*, vol. 42, no. 11, pp. 30-36, 1999.
- [17] A. Callahan and N. H. Shah, "Machine learning in healthcare," in *Key Advances in Clinical Informatics*, pp. 279-291, Academic Press, Cambridge, MA, USA, 2017.
- [18] L. Fu, "Knowledge discovery based on neural networks," *Communications of the ACM*, vol. 42, no. 11, pp. 47-50, 1999.
- [19] G. P. Zhang, "Neural networks for classification: a survey," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 30, no. 4, pp. 451-462, 2000.
- [20] W. Klossgen and J. M. Zytkow, *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, Oxford, UK, 2002.
- [21] R. Pari, M. Sandhya, and S. Sankar, "Risk factors based classification for accurate prediction of the Preterm Birth," in *Proceedings of the 2017 International Conference on Inventive Computing and Informatics (ICICI)*, pp. 394-399, IEEE, Coimbatore, India, November 2017.
- [22] "The global burden of preterm birth," *The Lancet*, vol. 374, 2009.
- [23] S. Beck, D. Wojdyla, L. Say et al., "The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity," *Bulletin of the World Health Organization*, vol. 88, no. 1, pp. 31-38, 2010.
- [24] S. Saigal and L. W. Doyle, "An overview of mortality and sequelae of preterm birth from infancy to adulthood," *The Lancet*, vol. 371, no. 9608, pp. 261-269, 2008.
- [25] I. Rudan, K. Y. Chan, J. S. Zhang et al., "Causes of deaths in children younger than 5 years in China in 2008," *The Lancet*, vol. 375, no. 9720, pp. 1083-1089, 2010.
- [26] M. C. Hogan, K. J. Foreman, M. Naghavi et al., "Maternal mortality for 181 countries, 1980-2008: a systematic analysis of progress towards Millennium Development Goal 5," *The Lancet*, vol. 375, no. 9726, pp. 1609-1623, 2010.
- [27] World Health Organization (WHO), *Commission on Information and Accountability for Women's and Children's Health. Keeping Promises, Measuring Results*, WHO, Geneva, Switzerland, 2015.
- [28] S. R. Walani and J. Biermann, "March of Dimes Foundation: leading the way to birth defects prevention," *Public Health Reviews*, vol. 38, no. 1, pp. 1-7, 2017.
- [29] Y. Dong and J.-L. Yu, "An overview of morbidity, mortality and long-term outcome of late preterm birth," *World Journal of Pediatrics*, vol. 7, no. 3, pp. 199-204, 2011.
- [30] J. R. G. Challis, S. J. Lye, W. Gibb, W. Whittle, F. Patel, and N. Alfaidy, "Understanding preterm labor," *Annals of the New York Academy of Sciences*, vol. 943, no. 1, pp. 225-234, 2001.
- [31] H. Witten Ian and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers, San Francisco, CA, USA, 2nd edition, 2005.
- [32] Y. Li, T. Li, and H. Liu, "Recent advances in feature selection and its applications," *Knowledge and Information Systems*, vol. 53, no. 3, pp. 551-577, 2017.
- [33] H. Huan Liu and L. Lei Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 491-502, 2005.
- [34] B. M. Mercer, R. L. Goldenberg, A. Das et al., "The preterm prediction study: a clinical risk assessment system," *American Journal of Obstetrics and Gynecology*, vol. 174, no. 6, pp. 1885-1895, 1996.
- [35] L. K. Goodwin, M. A. Iannacchione, W. E. Hammond, P. Crockett, S. Maher, and K. Schlitz, "Data mining methods find demographic predictors of preterm birth," *Nursing Research*, vol. 50, no. 6, pp. 340-345, 2001.
- [36] C. Gao, S. Osmundson, D. R. Velez Edwards, G. P. Jackson, B. A. Malin, and Y. Chen, "Deep learning predicts extreme preterm birth from electronic health records," *Journal of Biomedical Informatics*, vol. 100, p. 103334, 2019.
- [37] A. Weber, G. L. Darmstadt, S. Gruber et al., "Application of machine-learning to predict early spontaneous preterm birth among nulliparous non-Hispanic black and white women," *Annals of Epidemiology*, vol. 28, no. 11, pp. 783-789, 2018.

- [38] M. Mailath-Pokorny, S. Polterauer, M. Kohl et al., "Individualized assessment of preterm birth risk using two modified prediction models," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 186, pp. 42–48, 2015.
- [39] M. Son and E. S. Miller, "Predicting preterm birth: cervical length and fetal fibronectin," *Seminars in Perinatology*, vol. 41, no. 8, pp. 445–451, 2017.
- [40] U. Elaveyini, S. P. Devi, and K. S. Rao, "Neural networks prediction of preterm delivery with first trimester bleeding," *Archives of Gynecology and Obstetrics*, vol. 283, no. 5, pp. 971–979, 2011.
- [41] T. Khatibi, N. Kheyrikoochaksarayee, and M. M. Sepehri, "Analysis of big data for prediction of provider-initiated preterm birth and spontaneous premature deliveries and ranking the predictive features," *Archives of Gynecology and Obstetrics*, vol. 300, no. 6, pp. 1565–1582, 2019.
- [42] M. Colstrup, E. R. Mathiesen, P. Damm, D. M. Jensen, and L. Ringholm, "Pregnancy in women with type 1 diabetes: have the goals of St. Vincent declaration been met concerning foetal and neonatal complications?" *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 26, no. 17, pp. 1682–1686, 2013.
- [43] H. Blencowe, S. Cousens, M. Z. Oestergaard et al., "National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications," *The Lancet*, vol. 379, no. 9832, pp. 2162–2172, 2012.
- [44] L. Liu, H. L. Johnson, S. Cousens et al., "Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000," *The Lancet*, vol. 379, no. 9832, pp. 2151–2161, 2012.
- [45] M. A. Hall, "Correlation-based feature selection for discrete and numeric class machine learning," in *Proceedings of Seventeenth International Conference on Machine Learning*, pp. 359–366, Stanford, CA, USA, July 2000.
- [46] K. Y. Ngiam and I. W. Khor, "Big data and machine learning algorithms for health-care delivery," *The Lancet Oncology*, vol. 20, no. 5, pp. e262–e273, 2019.
- [47] J. G. Nam, S. Park, E. J. Hwang et al., "Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs," *Radiology*, vol. 290, no. 1, pp. 218–228, 2019.
- [48] Z. Ahmed, K. Mohamed, S. Zeeshan, and X. Dong, "Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine," *Database*, vol. 2020, 2020.
- [49] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, 2019.
- [50] S. Pereira, F. Portela, M. F. Santos, J. Machado, and A. Abelha, "Predicting preterm birth in maternity care by means of data mining," in *Portuguese Conference on Artificial Intelligence*, pp. 116–121, Springer, Berlin, Germany, 2015.
- [51] K. L. Courtney, S. Stewart, M. Popescu, and L. K. Goodwin, "Predictors of preterm birth in birth certificate data," *Studies in Health Technology and Informatics*, vol. 136, pp. 555–60, 2008.
- [52] N. S. Prema and M. P. Pushpalatha, "Machine learning approach for preterm birth prediction based on maternal chronic conditions," in *Emerging Research in Electronics, Computer Science and Technology*, pp. 581–588, Springer, Berlin, Germany, 2019.
- [53] F. Cunningham, K. Leveno, S. Bloom, C. Y. Spong, and J. Dashe, *Williams Obstetrics, 24e*, McGraw-Hill, New York, NY, USA, 2014.
- [54] R. Jin, Y. Breitbart, and C. Muoh, "Data discretization unification," *Knowledge and Information Systems*, vol. 19, no. 1, pp. 1–29, 2009.
- [55] B. K. Sarkar, S. S. Sana, and K. Chaudhuri, "MIL: a data discretisation approach," *International Journal of Data Mining, Modelling and Management*, vol. 3, no. 3, pp. 303–318, 2011.
- [56] G. Mitra, S. Sundareisan, and B. K. Sarkar, "A simple data discretizer," 2017, <http://arxiv.org/abs/1710.05091>.
- [57] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," *Computational and Structural Biotechnology Journal*, vol. 13, pp. 8–17, 2015.
- [58] N. Jothi, N. A. A. Rashid, and W. Husain, "Data mining in healthcare-a review," *Procedia Computer Science*, vol. 72, pp. 306–313, 2015.
- [59] M. Boulle, "Khiops: a statistical discretization method of continuous attributes," *Machine Learning*, vol. 55, no. 1, pp. 53–69, 2004.
- [60] P. N. Tan, M. Steinbach, and A. Karim, *Introduction to Data Mining*, Pearson Education India, New Delhi, India, 2016.
- [61] I. H. Witten and E. Frank, "Data mining," *ACM Sigmod Record*, vol. 31, no. 1, pp. 76–77, 2002.
- [62] A. Dhillon and A. Singh, "Machine learning in healthcare data analysis: a survey," *Journal of Biology and Today's World*, vol. 8, no. 6, pp. 1–10, 2019.
- [63] M. A. Ahmad, C. Eckert, and A. Teredesai, "Interpretable machine learning in healthcare," in *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 559–560, Washington, DC, USA, August 2018.
- [64] Q. Gu, L. Zhu, and Z. Cai, "Evaluation measures of the classification performance of imbalanced data sets," in *International Symposium on Intelligence Computation and Applications*, pp. 461–471, Springer, Berlin, Germany, 2009.
- [65] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
- [66] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," *NESUG Proceedings: Health Care and Life Sciences*, vol. 19, 2010.

Research Article

Big Data-Enabled Analysis of Factors Affecting Patient Waiting Time in the Nephrology Department of a Large Tertiary Hospital

Jialing Li ¹, Guiju Zhu ¹, Li Luo ², and Wenwu Shen ³

¹School of Management, Hunan University of Technology and Business, Changsha 410205, China

²Business School of Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu, China

³Outpatient Department, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

Correspondence should be addressed to Guiju Zhu; zhuguiju2010@163.com and Li Luo; luolicc@scu.edu.cn

Received 25 January 2021; Accepted 20 May 2021; Published 28 May 2021

Academic Editor: Giovanni Improta

Copyright © 2021 Jialing Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The length of waiting time has become an important indicator of the efficiency of medical services and the quality of medical care. Lengthy waiting times for patients will inevitably affect their mood and reduce satisfaction. For patients who are in urgent need of hospitalization, delayed admission often leads to exacerbation of the patient's condition and may threaten the patient's life. We gathered patients' information about outpatient visits and hospital admissions in the Nephrology Department of a large tertiary hospital in western China from January 1st, 2014, to December 31st, 2016, and we used big data-enabled analysis methods, including univariate analysis and multivariate linear regression models, to explore the factors affecting waiting time. We found that gender ($P = 0.048$), the day of issuing the admission card (Saturday, $P = 0.028$), the applied period for admission ($P < 0.001$), and the registration interval ($P < 0.001$) were positive influencing factors of patients' waiting time. Disease type (after kidney transplantation, $P < 0.001$), number of diagnoses ($P = 0.037$), and the day of issuing the admission card (Sunday, $P = 0.001$) were negative factors. A linear regression model built using these data performed well in the identification of factors affecting the waiting time of patients in the Nephrology Department. These results can be extended to other departments and could be valuable for improving patient satisfaction and hospital service quality by identifying the factors affecting waiting time.

1. Introduction

Long waiting times are recognized as a major obstacle to hospital care, affecting the quality of service and the establishment of friendly relationships with patients [1]. Due to the imbalance between the supply of and demand for medical resources, the problem of excessive waiting time is an issue for patients all over the world in outpatient, emergency, and hospitalization services. For patients, lengthy waiting time is not conducive to early treatment of their disease [2]. For hospitals, waiting time has an important influence on patient satisfaction [3]. Therefore, medical institutions committed to providing excellent service must effectively manage their clinic waiting times [4].

In some developed countries, the Emergency Department is the main channel for patient admission. For example, in the National University of Singapore Hospital,

emergency patients account for more than 64% of all inpatients [5]. In China, outpatient clinics are the main channel by which patients enter the hospital. In the West China Hospital (WCH) of Sichuan University, outpatients account for about 67% of inpatients. The huge increase in demand for hospitalization in China has brought significant challenges to the management of hospital beds in medical and health institutions. In particular, large public hospitals with good equipment and technical conditions often have long waiting times. Due to the shortage of hospital beds, some public hospitals in China, such as the West China Hospital of Sichuan University, have an average of more than 6000 patients waiting in line every day. The waiting time for hospitalization of patients averages three months to six months, and some waits are more than one year [6].

Both the Emergency Department and the Outpatient Department receive inpatients. Inpatients generally require

more services than outpatients, given the complexity of their diseases, so their waiting time is important. Therefore, in this study, we focused on factor analysis of the waiting time of admitted patients. To a certain extent, the hospitalization waiting time reflects the hospital's admission and discharge management level and the hospital's inpatient service quality [7]. Shortening the patient's admission waiting time will help improve the patient's admission experience and reduce the occurrence of medical disputes [8].

Different hospital departments have different admission characteristics, and the waiting time varies greatly in different departments. Therefore, under the advice of the administrator of the Admission Service Center, we chose the Department of Nephrology as the subject for the analysis of factors affecting the admission of patients, rather than discussing the determinants of inpatient waiting time for the entire hospital. This approach allows us to generate more specific, personalized admission management recommendations.

Several studies have demonstrated that long waiting times have negative impacts on the hospital's quality of service [9], patient satisfaction [10, 11], and hospital reputation [12]. Susanto and Chalidyanto [13] investigated waiting time and patient satisfaction in the pharmacy, using a cross-sectional study. Given the impact of waiting time on hospitals and patients, hospitals should take active measures to effectively manage the waiting times of patients.

There have been many studies into waiting time in healthcare institutions, focusing mainly on outpatients [14, 15] and Emergency Departments [16, 17]. The research has mostly been carried out from the perspective of the country, the region, or the entire hospital [18, 19]. For example, Geta and Edessa [20] investigated the factors affecting the waiting time of outpatients, using a questionnaire. Isfahani and colleagues [21] assessed the effects of a discharge lounge on decreasing patient waiting time and Emergency Department overcrowding, using a computer simulation. They found that the main factors leading to long patient waiting times are hindrances in patient flow and the occupation of Emergency Department beds by non-emergency patients. There have been relatively few studies into the waiting time of inpatients, and there is a lack of evidence-based management recommendations for hospitals.

Some research has focused on exploring which factors tend to increase the waiting time, including the overcrowding of patients, lack of healthcare providers, employee attitudes, work processes, length of hospital stay, and management problems [22–26]. However, this research primarily uses traditional methods, such as questionnaires, interviews [27], qualitative descriptions [28], and simulations [29]. For example, Aburayya et al. [30] collected questionnaires from 12 healthcare centers in the Emirate of Dubai in the UAE and found that the main causes of waiting time were high staff workloads, insufficient work procedures, employee-supervisor interaction problems, and lack of adequate facilities.

Data-driven methods are rarely used to explore the factors that affect the waiting times of inpatients. Data mining technology and machine learning methods have

been successfully applied in many fields, such as intelligent diagnosis and treatment [31–33], engineering [34], and security [35]. The wide range of these applications suggests that data mining technology may be used to analyze the factors that affect waiting time. Multivariate linear regression analysis, using statistical significance to identify explanatory variables, has been suggested as an effective method for evaluation, using big data [36].

In this study, we addressed this issue. We used the Department of Nephrology, WCH of Sichuan University, as an example, and analyzed the data from admitted patients, using multivariate linear regression, a machine learning algorithm, to unearth the key factors affecting the waiting time of inpatients. We used these data to provide evidence-based suggestions for reducing waiting times.

The remainder of this paper is structured as follows. Materials and methodology are introduced in Section 2. Results obtained using linear models are presented in Section 3. The analysis of the factors is presented in Section 4. In Section 5, we provide a brief conclusion.

2. Materials and Methods

2.1. Study Setting. West China Hospital is a large tertiary hospital in western China. It is faced with an admission problem common to large hospitals: bed resources cannot meet the demand for admission, and patients usually wait a long time before admission. In response to the increasing demand for hospitalization, an Admission Service Center was established in 2013 to centrally manage hospital beds. The data for this study came from the registration system of the Admission Service Center.

To better understand the waiting time of inpatients, we first provide a brief overview of the admission process for inpatients of WCH. Figure 1 is a schematic of the admission process.

2.1.1. Outpatient Service. First, each elective patient needs to see a doctor in the Outpatients Department. The outpatient doctors then provide a patient admission card, according to the severity of their illness. An admission card is an important certificate for hospitalization. Patients without an admission card will leave the hospital.

2.1.2. Admission Service Center. When a patient receives an admission card, they go to the Admission Service Center (ASC) to complete the hospital registration with information such as demographics, disease type, and insurance. After registration, a patient is added to the waiting list, sorted by registration date. A professional selects the patients who most need hospitalization services from the waiting list, based on their registration information. When a patient is selected, the professional will call the patient to ask if they have time to come to the hospital the next morning. Admitted patients who are notified by telephone and agree to hospitalization can go to the ASC to complete other procedures such as preoperative examinations, CT scans, and other diagnostic tests. The patient is conveyed to the ward at the appointed time.

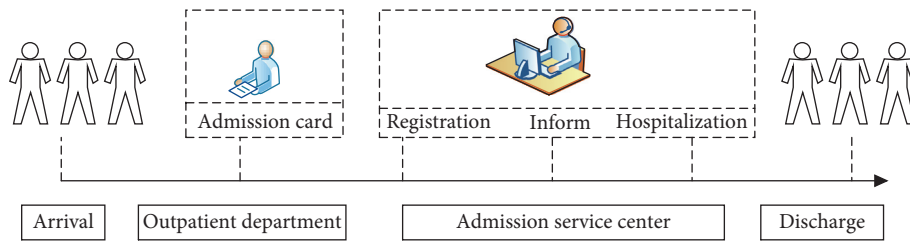


FIGURE 1: Flow chart of the elective patient admission process.

2.1.3. Discharge. Patients accept hospitalized services (including preoperative tests, treatment, and postoperative tests) and are finally discharged from the hospital after recovery.

Using this process, we defined the waiting time of inpatients as the time between the issuing of the patient's admission certificate by the outpatient doctor and the patient's formal admission for treatment.

2.2. Data Source. We first collected all of the data generated by the patients before admission from the Admission Service Center. The data included gender, age, date of application for admission (year, month, day, hour, and minute), registration date, outpatient diagnostic information, and subspecialty information. This information is the main resource used by medical staff when judging whether a patient is admitted to the hospital.

We extracted the admission data from the Department of Nephrology, West China Hospital, from January 1, 2014, to December 31, 2016. After deleting missing values and outliers, a total of 13,336 samples were obtained. All data were anonymized.

2.3. Data Preprocessing. After extracting the required fields, the data were preprocessed. We performed feature engineering on the original data to extract features for use in the model. The process of data processing is as follows.

First, males were coded as 1 and females as 2. Age was divided into four categories: juvenile (0–17 years old) as 1, young (18–40 years old) as 2, middle-aged (41–65 years old) as 3, and old (over 66 years old) as 4.

We then derived new fields based on the current fields. (1) Registration interval (RI) refers to the interval between the registration date and the date of issue of the admission card. (2) Standardized date of admission was split into two fields: the week of issuing the admission card (WIAC) and the applied period for admission (APA). We divided the latter into two periods: morning and afternoon. We labeled the weekday from 1 (Monday) to 7 (Sunday), and the applied period for admission was coded as 1 (morning) or 2 (afternoon). For example, a patient who had an admission card issued at 8 : 23 am on November 29, 2016, was coded as being issued with an admission card on Tuesday (assigned as 3) morning (assigned as 1). (3) Since many patients register on the same day after receiving the admission card, the week of the admission card date is very similar to that of the

registration date. In the end, only the week of the admission card date was retained. (4) The outpatient diagnosis information was the outpatient doctor's record of the patient's condition. A new field, the number of disease diagnoses, was derived by counting the number of diagnosed items.

The seven independent variables used in this study are shown in Table 1. We divided the following independent variables into three categories. The first category was descriptive statistical information, including gender and age. The second category was time information, which contains three variables: WIAC, APA, and RI. The third category was disease information, including two variables, TD and NDD. TD was divided into five subgroups: vascular access, renal biopsy, peritoneal dialysis, after kidney transplantation, and others. NDD was divided into four levels. In the outpatient diagnosis field, the diagnosis of only one disease was assigned a value of 1, diagnosis of two diseases was assigned 2, diagnosis of three diseases was assigned 3, and four or more diseases are assigned 4.

2.4. Methods. After sorting and converting the original data, R software was used for data analysis and modeling [37]. We first summarized the data. The measurement data was described by the mean, and the counting data was described by percentage.

We then carried out a univariate analysis. The difference in the waiting time of inpatients between each group was analyzed using univariate analysis [38], and the relationship between continuous variables and the waiting time of inpatients was analyzed using the Spearman correlation [39].

Finally, we constructed a multivariate linear regression model to explore the factors affecting the waiting time of patients registered in the Nephrology Department. We used the fields in Table 1 as independent variables and waiting time as the dependent variable. We used stepwise regression to filter the independent variables. Variables with $P < 0.05$ were selected as independent variables, at an inspection level of $\alpha = 0.05$. The generalized variance inflation factor (GVIF) and variance inflation factor were used to test the multivariate collinearity of the model. We assumed that there was no collinearity when GVIF or VIF was less than 2 [40].

3. Results

3.1. Descriptive Analysis. The descriptive analysis of the variables and sample information is shown in Figures 2–4. Figure 2 is a descriptive analysis diagram of the variables age

TABLE 1: Variable assignment table.

Categories	Variables	Assignment
Descriptive statistics	Age	1 = juvenile (0–17 years old), 2 = youth (18–40 years old), 3 = middle age (41–65 years old), 4 = old age (over 66 years old)
	Gender	1 = male, 2 = female
Time information	Week of issuing the admission card (WIAC)	1 = Monday, 2 = Tuesday, 3 = Wednesday, 4 = Thursday, 5 = Friday, 6 = Saturday, 7 = Sunday
	Applied period for admission (APA)	1 = morning (00:00–11:59), 2 = afternoon (12:00–23:59)
	Registration interval (RI)	Continuous variable
Disease information	Type of disease (TD)	1 = renal biopsy, 2 = peritoneal dialysis, 3 = vascular access, 4 = after renal transplant, 5 = other
	Number of disease diagnosis (NDD)	1 = 1 diagnosis, 2 = 2 diagnoses, 3 = 3 diagnoses, 4 = ≥4 diagnoses

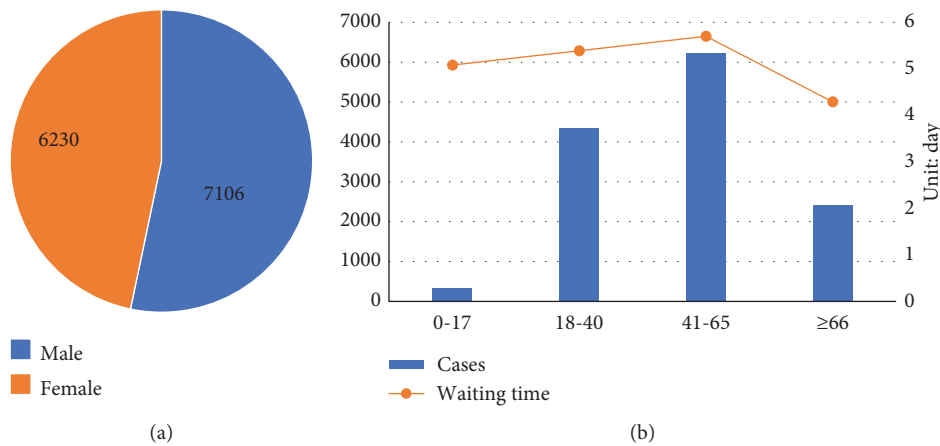


FIGURE 2: Characteristics of descriptive information. (a) Distribution of gender. (b) The number of cases by age and waiting time in days by age.

and gender. The dataset included 7,106 male patients, with an average waiting time of 5.72 days, and 6,230 female patients, with an average waiting time of 4.88 days. The number of male patients was slightly higher than that of female patients, and there was an approximately one-day gap in the average waiting time. The proportion of patients aged 41–65 was the largest, accounting for 46.75%. Middle-aged patients had the longest average waiting time, as high as 5.7 days.

Figure 3 shows an analysis of the time-related variables. During the admission period, we found that 56.46% of patients were registered in the morning (00:00–11:59), and 43.54% were registered in the afternoon (12:00–23:59). For the admission card, the number of patients on weekdays was much higher than that on weekends. Monday and Wednesday had the largest number of patients at 3,161 and 2,785, respectively. However, the waiting time on weekends was much higher than that on weekdays.

The characteristics of disease-related information are presented in Figure 4. In the Nephrology Department, TD is subdivided into five subspecialties. The proportion of cases of renal biopsy is the largest (33.05%), followed by vascular access (26.9%). The longest waiting time was for the peritoneal dialysis subspecialty (6.97 days). Most patients

(78.43%) had only one diagnosis, and the corresponding average waiting time was relatively long (5.34 days).

The average registration interval was 1.46 days. This observation indicates that some patients did not register for hospitalization immediately after receiving the admission card. The average waiting time for patients admitted to the Department of Nephrology was 5.33 days, and the standard deviation was 22.17 days. The high value of the standard deviation was due to the fact that some patients have milder disease, and the hospital always prioritizes the admission of severe cases, leaving some patients waiting for a long time.

We used univariate analysis and multivariate analysis to examine which factors affected the waiting time of inpatients, and how the data reflected the problems delaying the admission of patients.

3.2. Univariate Analysis. This section describes a univariate analysis of the factors affecting the waiting time of patients in the Nephrology Department. The results are shown in Table 2. Gender, week of issuing the admission card, and disease type had statistically significant effects on patient waiting time ($P < 0.05$). We then analyzed the correlation between the registration interval and the waiting time for admission. The waiting time of inpatients in the Nephrology

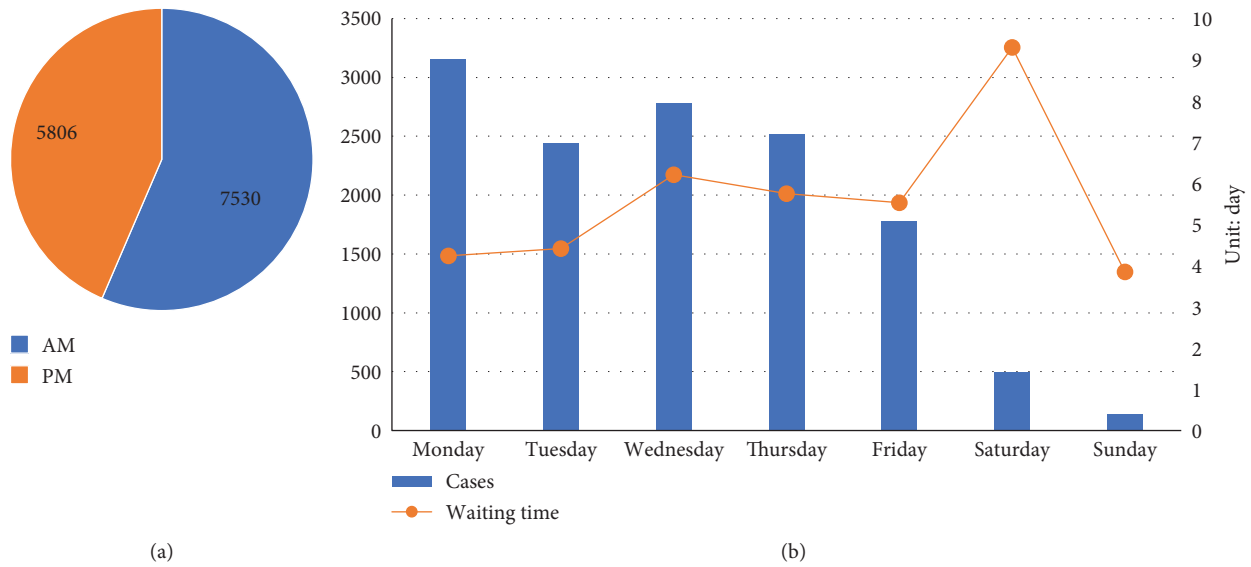


FIGURE 3: Characteristics of time-related information. (a) Pie chart of the applied period for admission. (b) Day of issuance of the admission card and waiting time by day.

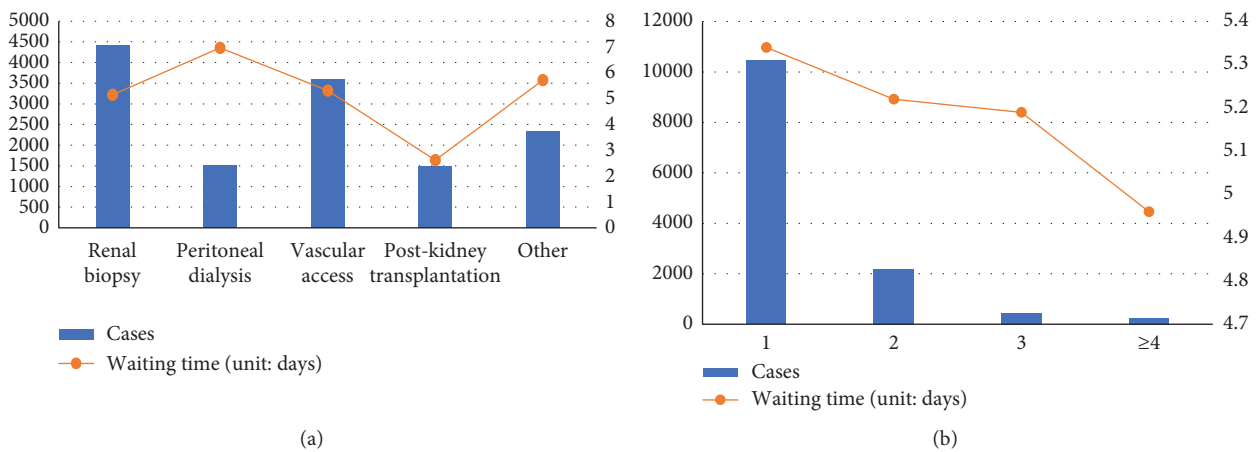


FIGURE 4: Characteristics of disease-related information. (a) Type of disease and waiting time. (b) The number of disease diagnoses and waiting time.

Department was positively correlated with registration interval ($P < 0.01$).

3.3. Multivariate Linear Regression Analysis. Taking the natural logarithm of waiting time as the dependent variable, we used stepwise regression to filter the independent variables. The week of admission, disease type, and the number of disease diagnoses were included in the model as dummy variables. The model results are as follows.

We used the coefficient of determination R^2 to measure the goodness of fit of the linear model. We saw an $R^2 = 0.527$, which shows that the regression line fits the observations well. For the model, we found $P < 0.001$, indicating that at a test level of $\alpha = 0.05$, the fit of the multivariate linear regression equation can be considered to be statistically significant.

The results of the multivariate linear regression analysis are shown in Table 3. We found that all variables except age were statistically significant.

Gender (male, $P = 0.048$), WIAC (Saturday, $P = 0.028$), APA (PM, $P < 0.001$), and RI ($P < 0.001$) were the positive factors influencing the waiting time. Taking the RI as an example, the unstandardized coefficient B_{was} was 0.789, and the 95% CI was (0.712, 0.736), indicating that when the other factors remained unchanged, for every additional day of the registration interval, the waiting time increased by 0.789 days.

The WIAC (Sunday, $P = 0.001$), DT (after renal transplant, $P < 0.001$), and NDD (Four or more diseases, $P = 0.037$) were negative factors. Taking the DT as an example, the coefficient B_{of} type four, after renal transplant was -3.091, which can be explained as follows: compared with type five, under the same conditions, the waiting time

TABLE 2: Comparison of the waiting time of hospitalized patients with different characteristics.

Variable	Cases	Waiting time	<i>F</i>	<i>P</i>
<i>Age</i>			2.365	0.069
Juvenile (0–17 years old)	341	5.08		
Youth (18–40 years old)	4 342	5.39		
Middle age (41–65 years old)	6 234	5.70		
Old age (over 66 years old)	2 419	4.29		
<i>Gender</i>			4.705	0.030
Male	7 106	5.72		
Female	6 230	4.88		
<i>WIAC</i>			5.648	<0.01
Monday	3 161	4.24		
Tuesday	2 443	4.42		
Wednesday	2 785	6.21		
Thursday	2 524	5.75		
Friday	1 784	5.53		
Saturday	496	9.30		
Sunday	143	3.85		
<i>APA</i>			2.365	0.124
Morning	7 530	5.12		
Afternoon	5 806	5.60		
<i>TD</i>			2.872	0.021
Renal biopsy	4 408	5.15		
Peritoneal dialysis	1 517	6.97		
Vascular access	3 595	5.31		
After renal transplant	1 492	2.62		
Other	2 324	5.72		
<i>NDD</i>			0.083	0.9
1 diagnosis	10 460	5.34		
2 diagnoses	2 190	5.22		
3 diagnoses	444	5.19		
≥4 diagnoses	242	4.96		

was reduced by 3.091 days for patients with the after renal transplant.

We performed a collinearity analysis on the above six statistically significant variables (Table 4). It can be seen that $GVIF < 2$ and $VIF < 2$, which means that there was no collinearity in the independent variables in our model.

The significance of the standardized regression coefficients β was to compare the importance of different independent variables to the dependent variable by standardizing B . The importance of each factor to the waiting time can be compared according to the absolute value of the standardized regression coefficient of linear regression. The results showed that RI (0.724), WIAC (Sunday, 0.191), DT (after renal transplant, 0.139), NDD (≥ 4 diagnoses, 0.097), and WIAC (Saturday, 0.077) were the top five factors that had the greatest impact on the waiting time for admission (Figure 5).

4. Discussion

This study was based on data provided by the information platform of the Admission Service Center and used data

mining technology to assess the determinants of inpatients' waiting time in the Nephrology Department of a tertiary hospital in western China. Based on the theoretical results presented in Section 3, we conducted interviews with the hospital's managers to learn more about the possible reasons behind the theoretical results. Finally, combined with observation of the actual situation of the hospital, we provide a theoretical basis for hospital administrators to take measures to shorten the waiting time for patients to be admitted to the hospital and improve the patient's admission experience.

4.1. Descriptive Information. This model showed that age had little effect on patients' admission, while gender was statistically significant. The waiting time of males admitted to the Department of Nephrology was 0.5 days longer than that of females. This is because there were more male than female outpatients (Table 2). Because Chinese hospitals require that men and women are not housed in the same ward, when the total number of beds is limited, higher numbers of outpatient visits for male patients lead to greater opportunity costs of waiting for beds, and the waiting time increases. This finding indicates that hospitals should adjust the ratio of male and female bed resource allocations and appropriately increase the number of admissions of male patients to meet different needs due to gender differences.

4.2. Time Information. The WIAC, RI, and APA had a considerable impact on the waiting time of patients. In general, patients who were issued an admission card on Sunday had shorter waiting times than those who were issued an admission card on Friday.

This result is related to the way in which the hospital is run and the resulting temporary release of hospital beds. The poor management of planned discharge from the clinical department of the hospital on weekends has led to the temporary discharge of some patients, and the Admissions Center has fewer staff on duty. After a patient is temporarily discharged, the hospital will admit patients who were issued an admission card and registered on-site on the same day. Therefore, patients who are issued an admission certificate on Sunday have shorter waiting times than those given a certificate on Friday.

From the perspective of the process of patient admission, the waiting time can be subdivided into indirect waiting time (the time between issuance of an admission card and registration at the Admission Center) and direct waiting time (the time between patients officially entering the waiting queue after registration and hospitalization). The registration interval reflects the urgency of admission, to a certain extent. The earlier a patient registers, the earlier he enters the hospital waiting system, and the shorter the direct waiting time. The length of the registration interval will directly affect the patient's waiting time for admission. Large hospitals should pay attention to the outpatient guidance service and guide patients to register effectively to reduce unnecessary waiting times.

Patients issued with admission permits in the afternoon wait longer. One possible reason is that the hospital releases fewer beds in the afternoon. According to the current

TABLE 3: Analysis of linear regression model of the waiting time and factors.

Variables	Unstandardized coefficient		Standardization coefficient β	T	P	95% confidence interval
	B	Standard error				
Intercept	5.647	1.087	0.066	1.353	0.1762	(-0.030, 0.162)
Gender (male)	0.527	0.266	0.024	1.978	0.048	(0.000, 0.047)
WIAC						
Thursday	0.042	0.473	0.002	0.088	0.930	(-0.040, 0.044)
Monday	-0.839	0.453	-0.038	-1.853	0.064	(-0.078, 0.002)
Sunday	-4.244	1.328	-0.191	-3.195	0.001	(-0.309, -0.074)
Saturday	1.704	0.777	0.077	2.194	0.028	(0.008, 0.146)
Wednesday	0.136	0.463	0.006	0.294	0.769	(-0.035, 0.047)
Tuesday	-0.777	0.476	-0.035	-1.633	0.103	(-0.077, 0.007)
APA (Afternoon)	1.057	0.268	0.048	3.941	0.000	(0.024, 0.071)
RI	0.789	0.007	0.724	121.349	0.000	(0.712, 0.736)
TD						
Renal biopsy	-3.091	0.741	-0.139	-4.172	0.000	(-0.205, -0.074)
Peritoneal dialysis	-0.520	0.351	-0.023	-1.481	0.139	(-0.054, 0.008)
Vascular access	-0.373	0.350	-0.017	-1.066	0.286	(-0.048, 0.014)
After renal transplant	1.212	0.722	0.055	1.679	0.093	(-0.009, 0.119)
NDD						
2 diagnoses	-1.445	0.993	-0.065	-1.455	0.146	(-0.154, 0.022)
3 diagnoses	-1.472	1.22	-0.065	-1.206	0.228	(-0.176, 0.040)
≥ 4 diagnoses	-2.156	1.035	-0.097	-2.082	0.037	(-0.190, -0.007)

Note. Bold fields indicate statistically significant variables.

TABLE 4: Collinearity diagnosis results of multivariate linear regression model.

Variables	GVIF	VIF
Gender	1.012	1.006
WIAC	1.036	1.003
APA	1.012	1.006
RI	1.002	1.001
TD	1.039	1.004
NDD	1.008	1.001

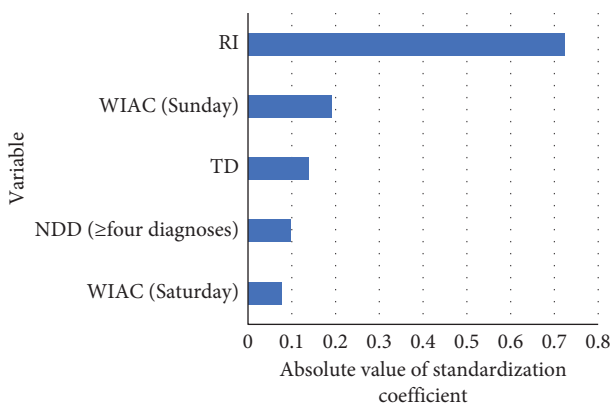


FIGURE 5: Primary important variable in the linear regression model.

operating rules of the hospital, the peak of admission and discharge is concentrated at 09:00–11:00 in the morning, lengthening the direct waiting time.

The hospital should strengthen the management of planned discharge. A reminder function can be added to the

prehospitalization information system to remind patients of the admission progress through means such as SMS alerts, or telephone follow-ups, to realize the whole-process tracking of patients' medical treatment, improve service quality, and improve the medical experience.

4.3. Disease Information. This study shows that disease type 4 (after renal transplant) had a shorter waiting time for admission than other disease types. This is because in patients with a after renal transplant, taking immunosuppressants leads to a decrease in immunity, resulting in rapid progression of infections and other complications, and an increased probability of systemic damage [41, 42]. Lengthy waiting times will seriously endanger their life, so the waiting time is shorter than those of other disease types.

Patients with more diagnostic items had shorter waiting times. This result reflects the impact of the severity of the disease on the waiting time, to a certain extent. The more diagnoses, the greater the probability that the disease is severe, and the shorter the waiting time compared to other, single disease, types. The results of this study on disease information are consistent with the hospital's actual practice. During the admission process of the case hospital used in this research, one of the rules is to focus on the type of disease and the severity of the disease, instead of scheduling on the principle of first come, first served. These findings suggest that in future admission management, the hospital should establish a more complete and detailed admission system centered on the type of disease, the severity of the disease, and the characteristics of the subspecialties in clinical departments to reduce the waiting time for critically ill patients.

The limitations and future research of this article can be summarized as follows. (1) The study does not consider the health system organizational model as a factor influencing the waiting times in the tertiary hospital. The coordination of primary care, hospitals, and communities improves early identification of health needs, healthcare service provision, and appropriateness [43, 44], reducing waiting times in hospitals and patient's satisfaction. (2) Further research is necessary to demonstrate that the indicators used are useful for the reorganization of services to reduce waiting times. In particular, it is necessary for healthcare professionals to improve the admission process based on the conclusions of this study to help further understand the factors that affect hospital admissions, such as disease care pathway, severity of disease, and services available in the territory [45–47]. (3) Taking into account the actual needs of the hospital, we will classify the waiting times and apply cutting-edge classification algorithms from the field of machine learning to accurately predict waiting times. Combining artificial intelligence technology with the needs of hospital admission will assist in hospitalization management and improve work efficiency.

5. Conclusions

The quality of medical service has become the core of hospital management, and continuous customer satisfaction should be the standard. The waiting time of patients in hospital is an important indicator and strongly affects patient satisfaction. In response to the need to reduce patient waiting time, we analyzed the status quo and found that although the average waiting time for admission in the Nephrology Department was less than one week, the standard deviation was large, indicating that the waiting time of patients at the individual level varies considerably, and there is still significant room for improvement.

It is important to find the key factors affecting waiting time for admission, to improve the status quo. Using a review of the literature, combined with interviews and clinical experience, we made full use of the historical data from the hospital, to identify the real problems reflected in the data, using data analysis. We constructed a general linear regression model and analyzed the factors that affected patients' waiting time for admission. The factors found in this study that have a significant impact on waiting time were gender (male), WAIC (Saturday and Sunday), APA (afternoon), RI, and DP (after renal transplant) and NDD (level 4).

The results of this research allow us to develop recommendations for hospital admission management, which can assist in patient management and improve work efficiency. We combined research interviews and literature analysis to provide suggestions for optimizing patients' admission times. These strategies can also reduce the psychological burden on patients.

Data Availability

The data used to support the findings of this study are restricted by the Admission Service Center of West China

Hospital in order to protect patient privacy. Data are available from West China Hospital for researchers who meet the criteria for access to confidential data.

Ethical Approval

This study protocol was approved by the Institutional Review Board (IRB) of West China Hospital in Sichuan University on the 21st of November 2019. The whole experiment complies with the ethical requirements and is completed under the relevant requirements.

Conflicts of Interest

The authors declare no conflicts of interest in this paper.

Acknowledgments

This project was sponsored by the National Office for Philosophy and Social Sciences Study on COST Control Optimization Model and PATH of Public Hospitals in China for DRGS Payment (19BGL248). The authors gratefully acknowledge the support from the Admission Service Center of West China Hospital of Sichuan University.

References

- [1] S. Al, N. M. A. Muhammad, and B. A. Kurdi, "Supply chain integration and customer relationship management in the airline logistics," *Theoretical Economics Letters*, vol. 9, no. 2, pp. 392–414, 2019.
- [2] C. Bleustein, D. B. Rothschild, A. Valen, E. Valatis, L. Schweitzer, and R. Jones, "Wait times, patient satisfaction scores, and the perception of care," *The American Journal of Managed Care*, vol. 20, no. 5, pp. 393–400, 2014.
- [3] G. M. Eilers, "Improving patient satisfaction with waiting time," *Journal of American College Health*, vol. 53, no. 1, pp. 41–48, 2004.
- [4] K. V. Rondeau, "Managing the clinic wait: an important quality of care challenge," *Journal of Nursing Care Quality*, vol. 13, no. 2, pp. 11–20, 1998.
- [5] P. Shi, "Models and insights for hospital inpatient operations: Time-dependent ED boarding time," *Management Science*, vol. 62, pp. 1–28, 2016.
- [6] T. Zhu, P. Liao, L. Luo, and H.-Q. Ye, "Data-driven models for capacity allocation of inpatient beds in a Chinese public hospital," *Computational and Mathematical Methods in Medicine*, vol. 202013 pages, 2020.
- [7] R. M. Al-dweeri, Z. M. Obeidat, M. A. Al-dwiry, M. T. Alshurideh, and A. M. Alhorani, "The impact of e-service quality and e-loyalty on online shopping: moderating effect of e-satisfaction and e-trust," *International Journal of Marketing Studies*, vol. 9, no. 2, pp. 92–103, 2017.
- [8] I. Umar, M. O. Oche, and A. S. Umar, "Patient waiting time in a tertiary health institution in Northern Nigeria," *Journal of Public Health and Epidemiology*, vol. 3, no. 2, pp. 78–82, 2011.
- [9] R. Clay-Williams, N. Taylor, H. P. Ting et al., "The relationships between quality management systems, safety culture and leadership and patient outcomes in Australian Emergency Departments," *International Journal for Quality in Health Care*, vol. 32, no. Supplement_1, pp. 43–51, 2020.
- [10] R. T. Anderson, F. T. Camacho, and R. Balkrishnan, "Willing to wait?: the influence of patient wait time on satisfaction with

- primary care,” *BMC Health Services Research*, vol. 7, p. 31, 2007.
- [11] S. Viotti, C. G. Cortese, J. Garlasco et al., “The buffering effect of humanity of care in the relationship between patient satisfaction and waiting time: a cross-sectional study in an emergency department,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, p. 2939, 2020.
- [12] D. I. Pillay, R. J. Ghazali, and N. H. Manaf, “Hospital waiting time: the forgotten premise of healthcare service delivery,” *International Journal of Health Care Quality Assurance*, vol. 7, pp. 506–522, 2011.
- [13] A. N. Susanto and D. Chalidyanto, “Waiting time and satisfaction of outpatient in the pharmacy section,” *EurAsian Journal of BioSciences*, vol. 14, pp. 3263–3266, 2020.
- [14] A. Aeenparast, S. J. Tabibi, K. Shahanaghi et al., “Reducing outpatient waiting time: a simulation modeling approach,” *Iranian Red Crescent Medical Journal*, vol. 15, p. 865, 2013.
- [15] F. P. Mardiah and M. H. Basri, “The analysis of appointment system to reduce outpatient waiting time at Indonesia’s public hospital,” *Human Resource Management Research*, vol. 3, pp. 27–33, 2013.
- [16] M. Ataman, “Predicting waiting and treatment times in emergency departments using ordinal logistic regression models,” *The American Journal of Emergency Medicine*, vol. 46, pp. 45–50, 2021.
- [17] K. Siamisang, J. T. Tlhakanelo, and B. B. Mhaladi, “Emergency department waiting times and determinants of prolonged length of stay in a Botswana referral hospital,” *Open Journal of Emergency Medicine*, vol. 8, no. 3, pp. 59–70, 2020.
- [18] I. Ibarra-Nava, V. Choudhry, and A. Agardh, “Desire to delay the first childbirth among young, married women in India: a cross-sectional study based on national survey data,” *BMC Public Health*, vol. 20, pp. 1–10, 2020.
- [19] J. Gaughan, P. Kasteridis, A. Mason, and A. Street, “Why are there long waits at English emergency departments?” *The European Journal of Health Economics*, vol. 21, no. 2, pp. 209–218, 2020.
- [20] E. T. Geta and A. M. Edessa, “Satisfaction with waiting time and associated factors among outpatients at nekemte referral hospital, western Ethiopia,” *Rehabilitation*, vol. 5, no. 2, pp. 18–25, 2020.
- [21] I. Mehdi, “Decreased emergency department overcrowding by discharge lounge: a computer simulation study,” *International Journal of Preventive Medicine*, vol. 11, 2020.
- [22] H. Babatabar-Darzi, I. Jafari-Iraqi, H. Mahmoudi et al., “Overcrowding management and patient safety: an application of the stabilization model,” *Iranian Journal of Nursing and Midwifery Research*, vol. 25, no. 5, p. 382, 2020.
- [23] A. Abu ELSamen and M. Alshuraideh, “The impact of Internal marketing on internal service quality: a case study in a Jordanian pharmaceutical company,” *International Journal of Business and Management*, vol. 7, no. 19, pp. 84–95, 2012.
- [24] Z. Alkalha, Z. Al-Zoubi, and M. Alshurideh, “Investigating the effects of human resources policies on organizational performance: empirical study on commercial banks operating in Jordan,” *European Journal of Economics, Finance and Administrative Sciences*, vol. 51, pp. 44–64, 2012.
- [25] A. Scala, A. M. Ponsiglione, I. Loperto et al., “Lean six sigma approach for reducing length of hospital stay for patients with femur fracture in a university hospital,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, p. 2843, 2021.
- [26] A. Shah, M. Memon, J. Kay, T. J. Wood, D. M. Tushinski, and V. Khanna, “Preoperative patient factors affecting length of stay following Total knee Arthroplasty: a systematic review and meta-analysis,” *The Journal of Arthroplasty*, vol. 34, no. 9, pp. 2124–2165, 2019.
- [27] G. Yadav, “A cross-sectional survey of population-wide wait times for patients seeking medical vs. cosmetic dermatologic care,” *PLoS One*, vol. 9, no. 11, p. e0162767, 2026.
- [28] A. M. Ponsiglione, “A Six Sigma DMAIC methodology as a support tool for Health Technology Assessment of two antibiotics,” *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 3469–3490, 2021.
- [29] G. Improta, G. Guizzi, C. Ricciardi et al., “Agile six sigma in healthcare: case study at Santobono pediatric hospital,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, p. 1052, 2020.
- [30] A. Aburayya, M. Alshurideh, A. Albqeen, D. Alawadhi, and I. A. A’yadeh, “An investigation of factors affecting patients waiting time in primary health care centers: an assessment study in Dubai,” *Management Science Letters*, vol. 10, pp. 1265–1276, 2020.
- [31] Li Luo, J. Li, C. Liu, and W. Shen, “Using machine learning methods to support healthcare professionals in making admission decisions,” *The International Journal of Health Planning and Management*, vol. 2, no. 34, pp. e1236–e1246, 2019.
- [32] Ayyoubzadeh and S. Mohammad, “A study of factors related to patients’ length of stay using data mining techniques in a general hospital in southern Iran,” *Health Information Science and Systems*, vol. 8, pp. 1–11, 2020.
- [33] S. Smys and S. Raj, “Internet of things and big data analytics for health care with cloud computing,” *Journal of Information Technology*, vol. 1, no. 1, pp. 9–18, 2019.
- [34] K. Yanamandra, G. L. Chen, X. Xu, G. Mac, and N. Gupta, “Reverse engineering of additive manufactured composite part by toolpath reconstruction using imaging and machine learning,” *Composites Science and Technology*, vol. 198, Article ID 108318, 2020.
- [35] B. T. Pham, A. Jaafari, M. Avand et al., “Performance evaluation of machine learning methods for forest fire modeling and prediction,” *Symmetry*, vol. 12, no. 6, p. 1022, 2020.
- [36] D. Qiao, Y. Zhang, A. u. Rehman, and M. R. Khosravi, “Big data-enabled analysis of DRGs-based payment on stroke patients in jiaozuo, China,” *Journal of Healthcare Engineering*, vol. 2020, Article ID 6690019, 9 pages, 2020.
- [37] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013.
- [38] V. Bewick, L. Cheek, and J. Ball, “Statistics review 9: one-way analysis of variance,” *Critical Care*, vol. 8, no. 2, p. 130, 2004.
- [39] P. Schober and T. R. Vetter, “Correlation analysis in medical research,” *Anesthesia & Analgesia*, vol. 130, no. 2, p. 332, 2020.
- [40] R. I. Kabacoff, *R in Action*, pp. 173–218, Manning publications, Shelter Island, NY, USA, 2011.
- [41] D. Palamuthusingam, K. Kunarajah, E. M. Pascoe et al., “Postoperative outcomes of kidney transplant recipients undergoing non-transplant-related elective surgery: a systematic review and meta-analysis,” *BMC Nephrology*, vol. 1, pp. 1–13, 2020.
- [42] M. Komagamine, T. Nishinaka, Y. Ichihara, S. Saito, and H. Niinami, “Long-term clinical outcomes of cardiac surgery for kidney transplant patients,” *Annals of Thoracic and Cardiovascular Surgery*, vol. 26, no. 2, pp. 84–87, 2020.

- [43] C. Tziraki, "Rethinking palliative care in a public health context: addressing the needs of persons with non-communicable chronic diseases," *Primary Health Care Research & Development*, vol. 21, 2020.
- [44] C. Tziraki-Segal, "Creating a culture of health in planning and implementing innovative strategies addressing non-communicable chronic diseases," *Frontiers in Sociology*, vol. 9, no. 4, pp. 1–15, 2019.
- [45] D. Sezgin, "The effectiveness of intermediate care including transitional care interventions on function, healthcare utilisation and costs: a scoping review," *European Geriatric Medicine*, vol. 11, no. 6, pp. 961–974, 2020.
- [46] D. Sezgin, A. Hendry, A. Liew et al., "Transitional palliative care interventions for older adults with advanced non-malignant diseases and frailty: a systematic review," *Journal of Integrated Care*, vol. 28, no. 4, pp. 387–403, 2020.
- [47] S. Lindner, "Can integrated care help in meeting the challenges posed on our health care systems by COVID-19? Some preliminary lessons learned from the European VIGOUR project," *International Journal of Integrated Care*, vol. 20, p. 4, 2020.

Research Article

A Hybrid Genetic Algorithm for Nurse Scheduling Problem considering the Fatigue Factor

Atefeh Amindoust ¹, Milad Asadpour ^{1,2} and Samineh Shirmohammadi³

¹Department of Industrial Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran

²Department of Information Systems and Operations Management, Business School, The University of Auckland, Auckland, New Zealand

³Young Researchers and Elite Club, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Correspondence should be addressed to Atefeh Amindoust; a.amindoost@pin.iaun.ac.ir

Received 20 January 2021; Revised 26 February 2021; Accepted 14 March 2021; Published 1 April 2021

Academic Editor: Giovanni Improta

Copyright © 2021 Atefeh Amindoust et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays and due to the pandemic of COVID-19, nurses are working under the highest pressure benevolently all over the world. This urgent situation can cause more fatigue for nurses who are responsible for taking care of COVID-19 patients 24 hours a day. Therefore, nurse scheduling should be modified with respect to this new situation. The purpose of the present research is to propose a new mathematical model for Nurse Scheduling Problem (NSP) considering the fatigue factor. To solve the proposed model, a hybrid Genetic Algorithm (GA) has been developed to provide a nurse schedule for all three shifts of a day. To validate the proposed approach, a randomly generated problem has been solved. In addition, to show the applicability of the proposed approach in real situations, the model has been solved for a real case study, a department in one of the hospitals in Esfahan, Iran, where COVID-19 patients are hospitalized. Consequently, a nurse schedule for May has been provided applying the proposed model, and the results approve its superiority in comparison with the manual schedule that is currently used in the department. To the best of our knowledge, it is the first study in which the proposed model takes the fatigue of nurses into account and provides a schedule based on it.

1. Introduction

In working places such as hospitals where the services must be provided continuously, distribution of workforces within different shifts is required. Accordingly, different studies could be found within the literature which formulated the problem of scheduling nurses, known as the Nurse Scheduling Problem (NSP), from different aspects to provide a timetable for nurses in a hospital [1–3].

In NSP, nurses are assigned different shifts according to a set of constraints and requirements which are determined by the hospital. On one hand, optimal nurse scheduling has effects on reducing hospital costs, increasing nurse's job satisfaction, quality of care, and increasing the input budget of hospitals [4]. On the other hand, hospital work

shifts could impose negative effects on involved nurses such as understaffing, heavy workload, and irregular work-scheduling conditions. It adversely affects the service quality of healthcare operations and leads to less patient-nurse interaction and patient safety issues. Meanwhile, other negative impacts of shift works on nurses are important such as fatigue, obesity, sleep disorder, and a wide range of chronic diseases [5–8]. It seems that human factors should be combined with NSP if a more productive hospital is needed. In other words, studying NSP considering human factors could improve both the performance of nurses and productivity of hospitals significantly [9].

In addition to former issues in nurses' work shifts, the pandemic of the COVID-19 has aggravated the problem and put double pressure on nurses across the world. In this new situation, work shifts should be determined regarding

response to patients' care requirements effectively and immediately and reduction of nurses' fatigue simultaneously.

Motivated by the aforementioned issues, and since, to the best of our knowledge, formulating an NSP considering fatigue factor has not studied yet, in this study, we have suggested a framework to distribute nurses in work shifts with the least fatigue alongside an attention to the additional pressure due to COVID-19 pandemic. To do so, a mathematical model is formulated for NSP in which fatigue of nurses has been taken into account. To solve the proposed model, a hybrid Genetic Algorithm (GA) has been developed, and real data from a department in one of the hospitals in Esfahan, Iran, where COVID-19 patients are hospitalized, has been used to provide a timetable for May.

The remaining parts of the current paper are as follows.

Firstly, related previous work on the area of NSP is investigated. After that, the problem is described and formulated, and then, the proposed GA is introduced. Following this, test problems are solved to approve the accuracy of the proposed model. Next, a real case study is considered, and a timetable for the department of COVID-19 patients is provided. Also, results are compared with the current timetable which is scheduled manually. Finally, the most important results along with suggestions for future research are presented in the conclusion section.

2. Literature Review

Taking preferences into account has been studied in NSP from different points of view. For instance, in the NSP study of Tsai and Li [10], Topaloglu and Selim [11], Abdollahi and Ansari [12], Wong et al. [13], Jafari et al. [14], and Legrain et al. [15], shift preferences were considered while in Vanhoucke and Maenhout [16] both shift preferences and day preferences were integrated simultaneously. Also, Liu et al. [17] considered NSP in hemodialysis service with a preference on roles and shifts. Further, Hamid et al. [9] proposed a multiobjective model for NSP considering the skill, preference, and compatibility of nurses.

In addition to preferences, some papers have added social/human factors into the NSPs. For example, in Farasat and Nikolaev [18], social structure effects into Nurse Scheduling Problem (NSP) were considered, and it was shown how the social structure of the working environment can affect the performance of nurses. In Zanda et al. [19], the maternity of nurses, type of their contracts (full time or part time), or nurses that benefit from special reductions of the workload for several reasons (e.g., because they have a close relative who suffers a serious illness) have been taken into account in providing a nurse schedule. Nahand et al. [20] proposed a multiobjective model for NSP considering human errors of nurses to determine optimal shift scheduling of nurses. Particularly, nurses' preference score, allocation costs, penalty cost of violating soft constraints, and human errors were all considered as objectives to be optimized, and a weighted-sum method was employed to solve the model. Wolbeck et al. [21] proposed a model for the NSP considering the timing and distribution of shift changes among nurses. In fact, the model incorporates a fair shift change

penalization scheme, in which the type, timing, and distribution of shift changes among nurses are taken into account. In addition, information from previous periods was considered, using an individual penalty score to distribute the shift changes fairly among the nurses. The model was solved within the Gurobi software.

Assigning nurses to the operating room is the subject of another group of papers in the NSP literature. Lim et al. [22] presented two models for NSP. The first one assigns nurses to upcoming surgery cases while the second one assigns lunch breaks for nurses. A column generation algorithm and a two-phase swapping heuristic were developed to find feasible assignments in a fast manner. In another study, Di Martinelly and Meskens [23] proposed a biobjective model for NSP to assign nurses to operating rooms. They have used the ϵ -constraint method to solve the problem. Besides, Chiang et al. [24] presented a mathematical model which can provide an optimized schedule for nurses and operating rooms simultaneously. GAMS software was deployed to solve the proposed model.

Also, categorizing patients into different groups and then assigning nurses to each group of them have been studied in NSPs as well. In particular, Heshmat et al. [25] studied a two-stage approach where, at the first stage, patients are clustered, and then, at the second stage, nurses are assigned to each cluster. The model was solved by using CPLEX. Moreover, Sarkar et al. [26] proposed a framework to assign nurses to home patients and patients in the hospital according to different demands with respect to the patients' health status. A hybrid GA was applied to solve the problem.

Meanwhile, the possibility of exchanging nurses within different departments has also been investigated. Fügener et al. [27] proposed a mid-term model for NSP considering cross-training effects. Firstly, a framework to define and visualize cross-training policies was proposed. Next, a new cross-training policy was introduced where each unit trained one dedicated nurse for each other unit, and finally, by using g CPLEX for solving the model, a schedule was provided. He et al. [28] proposed a two-stage stochastic model for NSP considering understaffing risk control. In the proposed model, an initial base number of nurses within the department's budget were assigned to the department. However, a central pool of nurses was maintained from where extra nurse shifts can be transferred from the pool to cover the shortage in certain departments, or redundant nurse shifts can be transferred into the pool. The model was solved using CPLEX. Schoenfelder et al. [29] presented a model that incorporates two classes of quick-response decisions in hospitals' nurse scheduling, namely, adjustments to the unit assignments of cross-trained float nurses and transfers of patients between units and off-unit admissions. They have examined the proposed model within different hospitals.

However, the main purpose of some papers is to develop heuristic/metaheuristic/simulation-based approaches for solving NSPs with the aim of improving the solution procedures from both quality and CPU time. In this case, hybrid Artificial Bee Colony (Awadallah et al. [30]), Particle Swarm Optimization (Wu et al. [31]), Sample Average Approximation method (Bagheri et al. [32]), Variable Neighborhood

Search algorithm (Zheng et al. [33], Rahimian et al. [34]), Hybrid Harmony Search algorithm (Awadallah et al. [35]), a heuristic approach based on Simulated Annealing (Knust and Xie [36]), a combination of three algorithms, namely, Fix-and-Relax, Fix-and-Optimize, and Simulated Annealing (Turhan and Bilgen [37]), and a heuristic procedure based on column generation (Strandmark et al. [38]) could be found within the literature of NSP.

According to a comprehensive literature review, we did not find any research which formulated NSP considering fatigue factor. To fill this research gap and to decrease double pressure which has been imposed on nurses due to the spreading of COVID-19, in the current paper, we formulated the NSP by considering the fatigue factor for a real case study, a department in one of the hospitals in Esfahan, Iran, where COVID-19 patients are hospitalized, to provide a schedule for nurses of this department for May in all shifts.

3. Problem Description

Health organizations such as hospitals are responsible for providing required services in both normal and crisis situations [39–41]. In recent years, health management has received special attention among researchers since appropriate allocation and usage of resources are vital to provide high-quality services to patients [42]. Accordingly, researchers in conjunction with decision-makers have worked on health management (as a pivotal issue in many political and social debates) in different regions and countries across the world to select an optimal form for the usage of healthcare resources [43]. To do so, they have applied novel decision-making tools to allocate and productive usage of resources [44] since the quality of the provided services to patients is highly dependent on the available resources. Among required resources, human resources are the most important factor in providing services [45, 46]. Because staff costs are a heavy part of each organization's cost, increasing the productivity and efficiency of human resources is very significant. One of the most practical ways to increase the productivity of this valuable resource is to combine the right distribution of human resources [47, 48]. In other words, it is very important to determine the number of required nurses during each shift so that, depending on the volume of arrivals, the provision of services always remains at the desired level. There are different types of shifts such as night shifts, last week shifts, two-part shifts, and on-call shifts. However, work shifts should be scheduled in such a way that imposes the lowest fatigue on nurses because long shifts have a negative efficacy on employee physiology. To do so, we propose a multiobjective model for the Nurse Scheduling Problem to optimize the work shifts of nurses, which includes three shifts in the morning, evening, and night. The first objective function is minimizing the total costs while the second objective function is minimizing the expected value of total fatigue for all nurses in all shifts. In fact, in this

model, since fatigue is one of the main human factors which affects the quality of doing jobs and health circumstances of employees, the nurse fatigue factor has been investigated as a function with binary values according to break times which makes a virtual life for nurses. In simple words, to reduce nurses' fatigue during the shifts, some certain times are allocated for their break.

Some of the assumptions of the problem are as follows:

- (i) All nurses have identical skills
- (ii) The break times are discrete values
- (iii) Demand behavior is the random variable based on a specific distribution function
- (iv) Each nurse is only assigned one shift
- (v) The break time is more important compared with shift time

3.1. Sets and Parameters.

- i : Period (day)
- j : Rest time (break time)
- c_1 : Employment cost
- c_2 : Dismissal cost
- c_3 : Nurse shortage cost
- c_4 : Nurse surplus cost
- c_5 : Fatigue reduction cost
- τ_{ij} : Duration of the rest time j at period i
- L_i : Duration of the shift time at period i
- Δ_i : Interval between the breaks at period i
- b : Maximum fatigue level
- a : positive value
- γ : positive value
- P : A very large number

3.2. Decision Variables and Functions.

- x_i : Nurse level changes at period i
- m : Rest level (integer positive values use for determining rest time)
- D : Required nurse number
- v_{ij} : Nursing attendance virtual time based on rest time j at period i
- $f(d)$: Probability function of required nurse number
- $F(t)$: Fatigue function
- $\delta(m)$: Fatigue fix function
- $c_{PM}(i)$: Nurses' fatigue reduction function at period i
- W : A binary variable, which is equal to 1 if nurse's employment is required; 0, otherwise
- Z : A binary variable, which is equal to 1 if nurse's dismissal is required; 0, otherwise

3.3. Model Formulation

$$\min f_1 = \sum_{i=1}^T c_{PM}(i) + c_1 \sum_{i=1}^T x_i^+ + c_2 \sum_{i=1}^T x_i^- + c_3 \sum_{i=1}^T \sum_{D=x_i}^{D_{\max}} (D - x_i) \frac{e^{-\lambda t} \cdot \lambda t^D}{D!} + c_4 \sum_{i=1}^T \sum_{D=x_i}^{D_{\max}} (x_i - D) \frac{e^{-\lambda t} \cdot \lambda t^D}{D!}, \quad (1)$$

$$\min f_2 = \sum_{i=1}^T x_i \left[\int_0^{\tau_{i1}} F(t) dt + \sum_{j=2}^n \int_{\tau_{ij-1}}^{\tau_{ij}} F(v_{ij} + t - \tau_{ij-1}) dt + \int_{\tau_{ij-1}}^{L_i} F(v_{in} + t - \tau_{in}) dt \right] \quad (2)$$

subject to

$$v_{ij} = v_{ij-1} + \delta(m) \cdot (\tau_{ij} - \tau_{ij-1}) \quad i = 1, \dots, T \quad j = 2, \dots, n, \quad (3)$$

$$x_{\min} \leq x_i \leq x_{\max} \quad i = 1, \dots, T, \quad (4)$$

$$x_i^+ \cdot x_i^- = 0 \quad i = 1, \dots, T, \quad (5)$$

$$x_i, x_i^+, x_i^-, D, v_{ij}, \geq 0 \quad m \geq 0, \text{ integer}, \quad (6)$$

$$W, Z \in \{0, 1\}. \quad (7)$$

Equations (1) and (2) are objective functions of the proposed model. The first objective function minimizes the total cost of nurses. The first sentence of this function relates to nurses' fatigue reduction function at period i . The second and third sentences of the first objective function calculate the costs related to decisions about the increase or reduction in the number of nurses in each shift. Finally, the fourth and fifth sentences of the first objective function determine the number of shortages and surplus of nurses at period i . Equation (2) is the second objective function and minimizes the expected value of total fatigue for all nurses in all shifts. Appendix A provides some complementary relations and explanations about these objective functions.

Also, constraints of the proposed model have been formulated as equations (3)–(7). Constraint (3) shows nursing attendance virtual time at period i . Constraint (4) determines a lower bound and upper bound for nurse level changes at period i . Constraint (5) assures that in each period solely either nurse employment or nurse dismissal could occur. Finally, constraints (6) and (7) define the domains of the decision variables.

3.4. Linearization. The above-proposed model is nonlinear because of constraint (5). Therefore, before proceeding to the next stage, constraint (5) should be substituted with identical linear equations. To do so, equations (8)–(10) should be used:

$$x_i^+ \leq P \cdot W, \quad (8)$$

$$x_i^- \leq P \cdot Z, \quad (9)$$

$$W + Z \leq 1. \quad (10)$$

4. Solution Approach

Genetic Algorithm (GA) is one of the first Evolutionary Algorithms in which selection, crossover, and mutation are its main operators. The idea behind the GA originates from the Darwinian Theory of Evolutionary and, therefore, GA is categorized as a population-based algorithm. It can be said that the GA is a programming technique that uses genetic evolution as a pattern for solving problems. The GA algorithm starts with a random population. This population contains a set of solutions, which represent chromosomes of individuals. The next step is creating the second generation of the population based on selection processes, i.e., generation based on the selected characteristics by genetic operators. For each individual, a pair of parents is selected. In the selection stage, the most appropriate elements will be selected so that even the weakest elements have the chance to choose. This process prevents from nearing the local answer.

Generally, GA has a connectivity probability that is between 0.6 and 1, which indicates the probability of birth of a child, and the organisms combine with this possibility. The connection of two chromosomes creates the child, which is added to the next generation. This process is done to find the right candidates for the answer in the next generation. This process creates a new generation of chromosomes that are different compared with the previous generation. The whole process is repeated until the last stage [49].

In comparison with other optimization algorithms, GA has certain advantages. The most important one is that GA has a powerful ability to tackle complex problems and could be used for solving a wide variety of optimization problems with different objective functions (e.g., stationary or nonstationary objective functions, linear or nonlinear objective functions, etc.). Moreover, since "multiple off-springs in a population act like independent agents, the population (or any subgroup) can explore the search space in many directions simultaneously. This feature makes it ideal to parallelize the algorithms for implementation while different parameters and even different groups of encoded strings can be manipulated at the same time" [50].

Due to the aforementioned advantages, in this study, a hybrid GA algorithm is applied to solve the proposed NSP model. Figure 1 shows the steps of the GA algorithm. Also, Appendix B provides more detail about the applied GA algorithm.

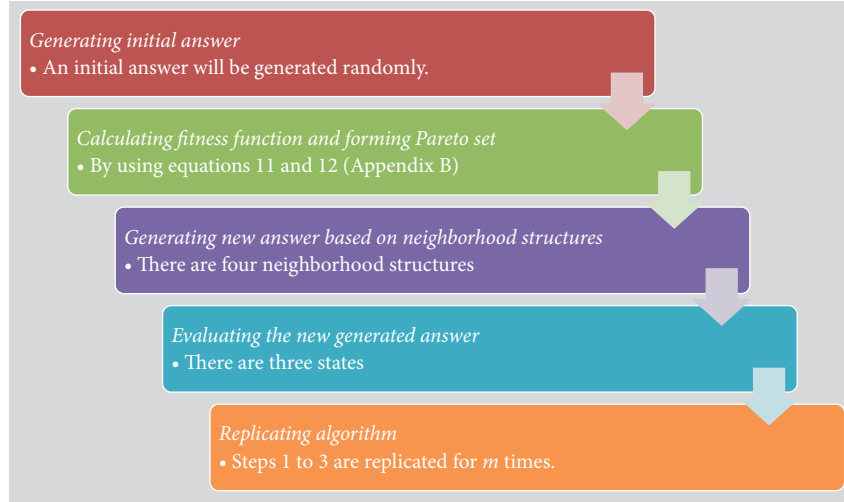


FIGURE 1: Steps of the GA algorithm.

5. Computational Results

5.1. Model Validation. Prior to solving the model by using real data, to validate the accuracy of the proposed model, a test problem is generated randomly and a schedule for one week is provided. Table 1 includes parameters for the test problem. All computations have been done within MATLAB software on a laptop with Intel Core i5, 2.5 GHz, and 4 GB of RAM.

The number of nurses is assumed to be five, and there are three shifts per day. Each nurse's salary per hour is also assumed to be 10 units per hour.

Table 2 shows the obtained schedule for a week for the test problem.

In Table 2, columns x_i^+ and x_i^- show the required changes in the number of nurses so that the value in column D is affected by these changes in each row. Also, column F1 contains the imposed costs to the system by implementing the schedule whereas column F2 shows the fatigue of nurses.

5.2. Case Study. After validating the proposed model, we performed the model in a real case, a specific department of a hospital in Esfahan, Iran, where COVID-19 patients are hospitalized. Our purpose was to compare the scheduling obtained using the proposed model with those currently used in the considered department which have been computed manually by a person in the human resource department of the case study. Due to the nature of constraints in the NSP of the case study, the person who is responsible for providing the schedule, firstly, arranges an initial schedule that distributes all nurses within the days of a month. After that, she/he tries to rearrange the scheduling in order to also meet the fatigue of nurses as well. Accordingly, administrators of the department noted that the manual approach could be very time-consuming. In addition, some nurses were enjoying generous rest hours while most nurses were under high pressure. Thereby, administrators decided to automate the scheduling of this department. It led to the

TABLE 1: Input parameters of test problem.

Parameters	Values
c_1, c_2, c_3, c_4	Uniform (1, 10)
λ	1
B	1
θ	20

TABLE 2: Schedule for the test problem.

Day (i)	x_i^+	x_i^-	m	D	F1 (cost)	F2 (fatigue)
1	0 0 0	0 0 0	0 0 0	2 1 1	2,156	0.0386
2	0 0 1	0 0 0	0 0 1	2 1 2	2,798	0.0369
3	0 1 0	0 0 1	2 1 0	2 2 1	2,604	0.0358
4	0 0 0	0 1 0	2 0 0	2 1 1	2,309	0.0258
5	0 0 0	0 0 0	1 0 0	2 1 1	2,255	0.0344
6	0 0 1	0 0 0	1 0 1	2 1 2	2,885	0.0361
7	0 1 0	0 0 1	1 1 0	2 2 1	2,515	0.029

distribution of nurses within the shifts in a fair manner while the required time for providing a schedule was reduced significantly as well.

The department has 14 nurses, and working days are partitioned into three shifts: 8: 00 am-4: 00 pm, 4: 00 pm-12: 00 pm (0: 00 am), and 12: 00 pm-8: 00 am. Each nurse's salary per hour is 160000 IRR. It is worth mentioning that 1 EUR is equal to almost 46,292 IRR. Based on the salary, costs of employment, dismissal, shortage and, surplus are as follows:

$$\begin{aligned}
 c_1 &= 7 * 8 * \text{nurse's salary per hour,} \\
 c_2 &= 22 * 8 * \text{nurse's salary per hour,} \\
 c_3 &= L_i * \text{nurse's salary per hour,} \\
 c_4 &= 2 * L_i * \text{nurse's salary per hour.}
 \end{aligned} \tag{11}$$

The patient's arrival rate is based on the Poisson distribution and is $\lambda_1 = 2$, $\lambda_2 = 3$, and $\lambda_3 = 3$ in the first, second, and third shifts. The fatigue factor follows the Weibull distribution with parameters $\beta = 2$ and $\theta = 500$. Also,

TABLE 3: Scheduling of May provided by hybrid GA and manually.

Day (i)	Manually (hospital)						Hybrid GA					
	x_i^+	x_i^-	m	D	F1 (cost)	F2 (fatigue)	x_i^+	x_i^-	m	D	F1 (cost)	F2 (fatigue)
1	0 0 0	0 0 0	0 0 0	6 3 3	71,481	0.0591	0 0 0	0 0 0	0 0 0	6 3 3	71,481	0.0591
2	0 0 0	0 0 0	0 0 1	6 3 3	73,123	0.0574	0 0 0	0 0 0	0 0 1	6 3 3	73,123	0.0574
3	0 0 0	0 0 0	0 0 2	6 3 3	76,729	0.0563	0 0 0	0 0 0	0 0 2	6 3 3	76,729	0.0563
4	0 0 0	0 0 0	1 0 0	6 3 3	74,992	0.0581	0 1 0	1 0 0	2 0 2	5 4 3	75,834	0.0463
5	0 0 0	0 0 0	1 0 1	6 3 3	77,440	0.0549	1 0 0	0 1 0	1 0 1	6 3 3	77,440	0.0549
6	0 0 0	0 0 0	1 0 2	6 3 3	78,070	0.0535	0 0 0	1 0 0	2 0 2	5 3 3	73,810	0.0566
7	0 0 0	0 0 0	1 1 1	6 3 3	81,770	0.0532	0 0 0	0 0 1	2 0 2	5 3 2	69,410	0.0495
8	0 0 0	0 0 0	1 1 2	6 3 3	89,200	0.0498	0 0 0	0 0 0	2 1 2	5 3 2	71,618	0.0448
9	0 0 0	0 0 0	2 0 1	6 3 3	79,250	0.0541	1 0 1	0 0 0	2 0 1	6 3 3	79,250	0.0541
10	0 0 0	0 0 0	2 0 2	6 3 3	83,880	0.0526	0 0 0	1 0 1	2 2 2	5 3 2	72,431	0.0507
11	0 0 0	0 0 0	1 1 2	6 3 3	89,200	0.0498	0 1 1	0 0 0	2 1 2	5 4 3	91,012	0.0455
12	0 0 0	0 0 0	2 2 2	6 3 3	92,310	0.0479	1 0 0	0 1 0	2 2 2	6 3 3	92,310	0.0479
13	0 0 0	0 0 0	1 0 2	6 3 3	78,070	0.0535	0 0 0	0 0 0	1 0 2	6 3 3	78,070	0.0535
14	0 0 0	0 0 0	2 2 2	6 3 3	92,310	0.0479	0 0 0	0 0 0	2 2 2	6 3 3	92,310	0.0479
15	0 0 0	0 0 0	2 0 0	6 3 3	74,235	0.0588	0 0 0	0 0 0	2 0 0	6 3 3	74,235	0.0588
16	0 0 0	0 0 0	1 0 1	6 3 3	77,440	0.0549	0 0 0	1 0 0	2 1 1	5 3 3	74,142	0.0571
17	0 0 0	0 0 0	2 0 2	6 3 3	83,880	0.0526	1 0 0	0 0 0	2 0 2	6 3 3	83,880	0.0526
18	0 0 0	0 0 1	2 0 2	6 3 2	68,300	0.0598	0 0 0	0 0 1	2 0 2	6 3 2	68,300	0.0598
19	0 0 0	0 0 0	2 1 2	6 3 2	70,173	0.0576	0 0 0	0 0 0	2 1 2	6 3 2	70,173	0.0576
20	0 0 0	0 0 0	2 2 2	6 3 2	72,345	0.0561	0 0 0	0 0 0	2 2 2	6 3 2	72,345	0.0561
21	0 0 0	0 0 1	2 0 2	6 3 1	60,900	0.0661	0 0 0	1 0 0	2 2 2	5 3 2	63,476	0.0609
22	0 0 0	0 0 0	1 1 2	6 3 1	62,200	0.0649	0 0 0	0 0 0	2 2 0	5 3 2	60,150	0.0637
23	0 0 0	0 0 0	1 1 0	6 3 1	59,145	0.0674	0 1 1	0 0 0	2 1 1	5 4 3	63,182	0.0528
24	0 0 0	0 1 0	1 1 1	6 2 1	51,165	0.0801	0 0 0	1 2 1	2 2 2	4 2 2	52,750	0.0746
25	0 1 0	1 0 0	2 2 2	5 3 1	54,100	0.0681	1 1 0	0 0 1	2 2 2	5 3 1	54,100	0.0681
26	0 0 1	0 1 0	2 2 2	5 2 2	56,108	0.0768	0 0 1	0 1 0	2 2 2	5 2 2	56,108	0.0768
27	1 1 0	0 0 0	2 2 2	6 3 2	72,345	0.0561	0 1 1	0 0 0	1 1 1	5 3 3	71,230	0.0574
28	0 0 0	2 1 0	2 2 2	4 2 2	52,215	0.0795	1 0 0	0 1 1	2 2 2	6 2 2	58,900	0.0646
29	0 0 1	0 0 0	2 2 2	4 2 3	54,000	0.0835	0 1 1	0 0 0	2 1 2	6 3 3	71,700	0.0513
30	1 1 0	0 0 2	2 2 2	5 3 1	54,100	0.0681	0 0 0	1 1 0	2 1 2	5 2 3	55,700	0.0662
31	0 0 1	1 1 0	0 0 0	4 2 2	51,985	0.0919	1 0 0	0 0 0	2 1 2	6 2 3	69,150	0.0598

the parameters of the hybrid GA algorithm including the percentage of mutation, intersection percentage, and the number of repetitions are 0.1, 0.9, and 200, respectively. The model is solved within MATLAB software on a laptop with Intel Core i5, 2.5 GHz, and 4 GB of RAM.

The schedule for May which has been prepared manually and the schedule that has been prepared by hybrid GA have been compared in Table 3. In Table 3, column *D* depicts the order of each shift in each day. For instance, 6 3 3 on the 1st day of May states that, on the first day of May, the number of required nurses for the first, second, and third shifts is six, three, and three, respectively. In addition, columns x_i^+ and x_i^- depict the level of changes required for each day regarding the last day. For example, on the 25th of May, one nurse should be added for a second shift whereas one nurse should be eliminated in the first shift compared with the 24th of May. Therefore, the order of shifts on the 25th of May is 5 3 1 while on the 24th of May the order was 6 2 1.

Column *m* presents the rest level at each shift. The value of 0 means that nurses will not be allowed to have further rest during their shift. In fact, all nurses have a break time generally in their shifts. If this usual break is the only break during the shift, *m* will give a value of 0. However, values of 1 and 2 for *m* state that nurses have one or two further breaks

during their shift. Undoubtedly, the surplus break times reduce the fatigue of nurses dramatically.

Finally, columns F1 and F2 contain the value of costs and fatigue function in each day. It should be mentioned that F1 values have been divided into 1,000 for more convenience.

Figure 2 compares costs of manual scheduling and scheduling calculated by hybrid GA algorithm for all 31 days in May.

As Figure 2 illustrates, the costs of the department during one month according to manual scheduling are lower than those of the hybrid GA algorithm. However, this difference is not very significant, and Figure 3 clearly approves this matter. As it can be seen in Figure 3, the total costs of manual scheduling are 2,212,461,000 IRR while the total costs of scheduling which have been prepared by the proposed model and using the hybrid GA algorithm are 2,214,349,000 IRR. Therefore, the difference between costs is not higher than 2,000,000 IRR, and actually, this is not a considerable amount of money.

Figure 4 compares fatigue function in manual scheduling and hybrid GA scheduling.

Figure 4 clearly reveals that although the values of fatigue function in the second decade of May are almost identical in both hybrid GA scheduling and manual scheduling, the

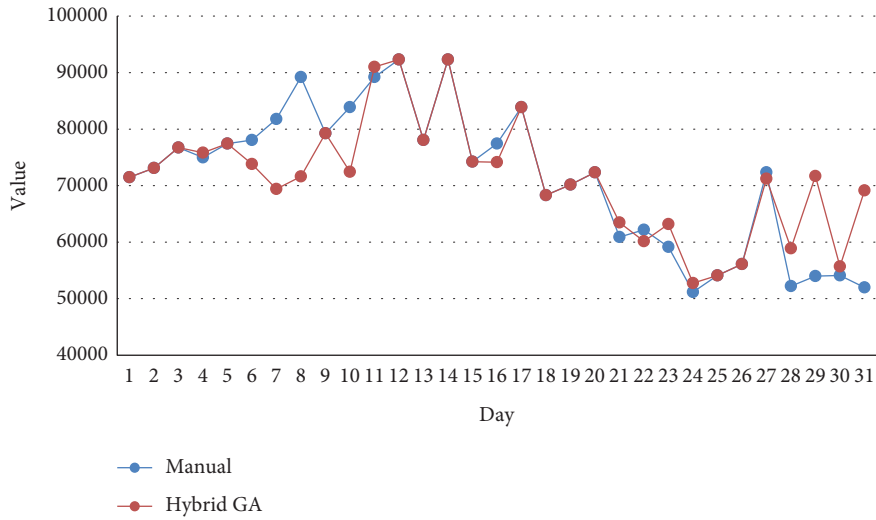


FIGURE 2: Comparison of the cost function.

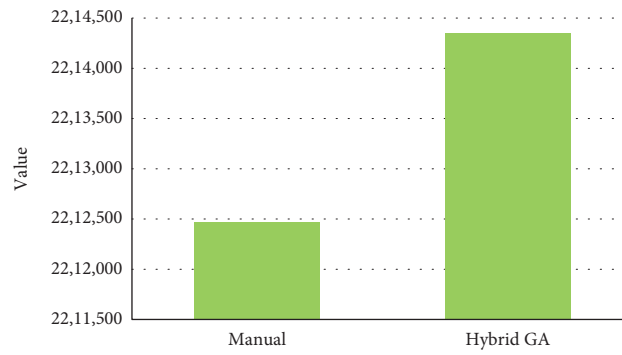


FIGURE 2: Cumulative values of the cost function.

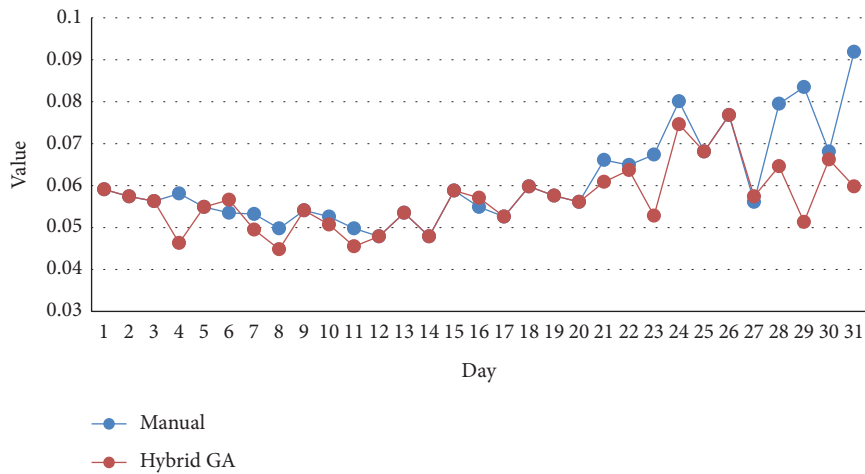


FIGURE 4: Comparison of fatigue function.

fatigue of nurses in hybrid GA scheduling is totally lower than that in manual scheduling. Figure 5 which shows cumulative values of fatigue function for 31 days of May obviously approves this claim.

In addition, the required total time for providing a timetable by hybrid GA is considerably lower than the required total time for providing a timetable manually. Figure 6 depicts this difference.

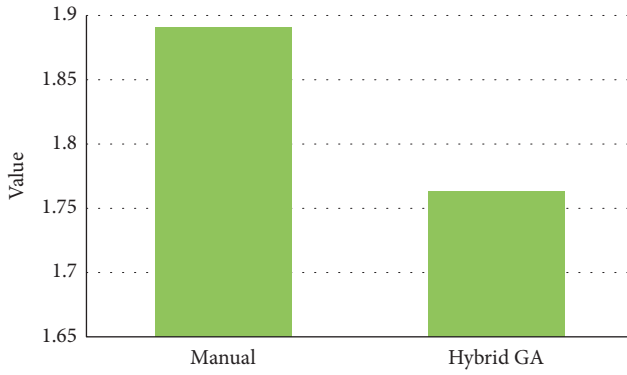


FIGURE 5: Cumulative values of fatigue function.

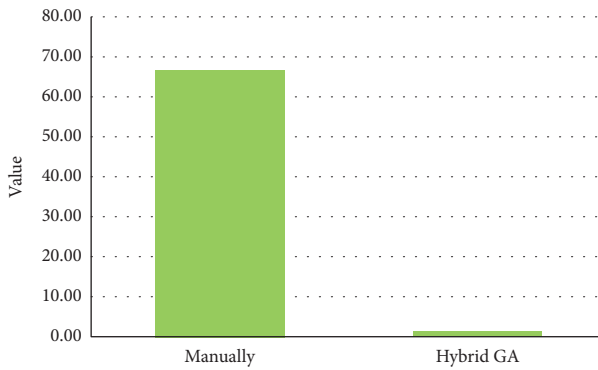


FIGURE 6: Required time for providing a schedule.

Figure 6 illustrates that scheduling of the department manually takes 66.67 minutes whereas the hybrid GA algorithm reduces the required time for scheduling dramatically so that it needs just 1.28 minutes for solving the proposed model and providing a timetable for the department.

6. Conclusion

Scheduling work shifts in many industry and service occupations directly affects the mental and physical health of employees. This matter can be more obvious in occupations like nursing where employees are usually working under high pressure and stress. Therefore, it would be better to consider employees' satisfaction and convenience in providing scheduling so that work hours impose lower pressure on them.

During the last months and due to the pandemic of COVID-19, physicians, nurses, and other persons in hospitals across the world are working under the highest pressure benevolently to suppress this insidious disease. This urgent situation can cause more fatigue for these persons, especially for nurses who are caring for COVID-19 patients 24 hours a day. Therefore, nurse scheduling should be modified with respect to this new situation.

In this paper, a mathematical model has been proposed to optimize the work shift of nurses considering human factors. To do so, according to experts' points of view, fatigue as the most important factor among human factors has been

selected, and minimizing fatigue by creating different time intervals of rest during shift works has been addressed.

To solve the proposed model, a hybrid GA algorithm has been performed. In addition, to validate the efficiency of the proposed model, real data from the department of COVID-19 patients in one of the hospitals in Esfahan have been used where scheduling is currently determined manually. In fact, applying the proposed procedure of this paper, a timetable for all shifts of this department has been provided for May, and results have been compared with those currently used in the considered department and provided manually. Results approve that while the proposed approach imposes a little more cost to the department, it outperforms the current manual scheduling in both fatigue factor and time of providing a time table.

Apart from the lower required time for providing a timetable compared with the former manual approach, one of the main advantages of the proposed approach is eliminating personal tastes. It has been done because in the proposed approach, no one is involved in providing schedules, and the level of available resources is the only effective factor on the final schedule. To sum up, the proposed approach distributes nurses within the shifts in a fair manner considering the least fatigue while the required time for providing a schedule is significantly lower than the manual approach as well. However, another strength of our proposed approach is its generalizability. In other words, due to the structure of formulation, the proposed model could be used in either other healthcare departments or other continuous systems where a schedule for different shifts is required. In all of these settings, the model could distribute nurses/workforces within different shifts fairly considering least fatigue.

In future research, our model can be developed by considering other human factors such as rewards, motivation, and loyalty. Also, developing other heuristic and metaheuristic techniques would be an interesting direction for future research.

Appendix

A. Complementary Relations

As mentioned in Section 3.3, the first objective function minimizes the total cost of nurses. The first sentence of this function relates to nurses' fatigue reduction function at period i which is calculated based on equation (A.1):

$$c_{PM}(i) = m \cdot \Delta_i \cdot c_5 \cdot a^{m \cdot \gamma}, \quad i = 1, \dots, T, \quad (\text{A.1})$$

The second and third sentences of the first objective function calculate the costs related to decisions about the increase or reduction in the number of nurses in each shift. In other words, x_i^+ depicts the number of employed nurses while x_i^- shows the number of dismissed nurses at period i . Finally, the fourth and fifth sentences of the first objective function determine the number of shortages and surplus of nurses at period i . On the one hand, when demand is higher than the number of nurses ($D > x_i$), the cost of shortage could be computed as follows:

$$c_3 \sum_{i=1}^T (D - x_i) f(d). \quad (\text{A.2})$$

On the other hand, when demand is less than the number of nurses ($D \leq x_i$), the cost of the surplus could be computed as follows:

$$c_4 \sum_{i=1}^T (D - x_i) f(d) \quad D \leq x_i. \quad (\text{A.3})$$

It should be noted that x_i is computed based on equation (A.4) and (A.5):

$$x_i = x_{i-1} + x_i^+ - x_i^- \quad i = 1, \dots, T. \quad (\text{A.4})$$

Since D is a Poisson random variable, $f(d)$ in equations (A.2) and (A.3) is a Poisson distribution function. Therefore,

equations (A.2) and (A.3) should be rewritten as the following equations:

$$c_3 \sum_{i=1}^T (D - x_i) \frac{e^{-\lambda t} \cdot \lambda t^D}{D!} \quad D > x_i, \quad (\text{A.5})$$

$$c_4 \sum_{i=1}^T (D - x_i) \frac{e^{-\lambda t} \cdot \lambda t^D}{D!} \quad D \leq x_i.$$

Moreover, as mentioned in Section 3.3, the proposed model has a second objective function which minimizes the expected value of total fatigue for all nurses in all shifts. In the current objective function, there are some values that will be calculated using the following equations:

$$v_{ij}(t) = \begin{cases} t_i & t_i < \tau_{i1} & i = 1, \dots, T \\ v_{ij-1} + t + \tau_{ij-1} & \tau_{ij-1} \leq t < \tau_{ij} & j = 2, \dots, n' \end{cases}$$

$$F(v_{ij}(t)) = \begin{cases} F(t) & t_i < \tau_{i1}, F(t) < b & i = 1, \dots, T \\ F(v_{ij-1} + t - \tau_{ij-1}) & \tau_{ij-1} \leq t < \tau_{ij}, F(t) < b & i = 1, \dots, T, j = 2, \dots, n, \\ 0 & F(t) \geq b. \end{cases} \quad (\text{A.6})$$

B. Steps of the Hybrid GA Algorithm

1) *Generating Initial Answer.* Since there are four variables including x_i^+ , x_i^- , m , D and three shifts, all feasible solutions will be revealed in the form of a 4×3 matrix. A feasible answer can be found as follows:

$$\begin{array}{lcl} x_i^+ & \longrightarrow & 0 \quad 1 \quad 0 \\ x_i^- & \longrightarrow & 2 \quad 0 \quad 1 \\ m & \longrightarrow & 1 \quad 2 \quad 1 \\ D & \longrightarrow & 4 \quad 4 \quad 2 \end{array} \quad (\text{B.1})$$

An initial answer will be generated randomly. For this purpose, one of the first or second rows is selected randomly. To fill the first column, a random number between $[x_{\min}, x_{\max}]$ is selected. If the selected number is greater than zero, we will write a zero in the next row while if the selected number is zero, a random number in the interval $[x_{\min}, x_{\max}]$ will be selected for another row. The same procedure will be employed for the second and third columns as well. In the third row, a random number in the interval $[x_{\min}, x_{\max}]$ opts for the first to third columns respectively.

2) *Fitness Function.* Suppose $f_1 = c_1 \sum_{i=1}^T x_i^+$ and $f_2 = c_2 \sum_{i=1}^T x_i^-$. The fitness function is calculated according to these functions and will put in the Pareto set.

3) *Generating New Answer Based on Different Neighborhood Structures.* The first neighborhood structure is as follows:

Between the first and second rows, one of them is chosen randomly. Suppose that i represents the selected row. Now a random column like j is chosen randomly. If $x_{ij} + 1 \leq x_{\max}$, then $x_{ij} + 1 \longrightarrow x_{ij}$ and the value of column j in the next row is zero.

The second neighborhood structure:

Between the first and second rows, one of them is chosen randomly. Suppose that i represents the selected row. Now, a random column like j is chosen randomly. If $x_{ij} - 1 \geq x_{\min}$, then $x_{ij} - 1 \longrightarrow x_{ij}$.

The third neighborhood structure:

In the third row, a column j is chosen randomly. If $m_j + 1 \leq m_{\max}$, then $m_j + 1 \longrightarrow m_j$.

The fourth neighborhood structure:

In the third row, a column j is chosen randomly. If $m_j - 1 \geq m_{\min}$, then $m_j - 1 \longrightarrow m_j$.

The algorithm starts with the first neighborhood structure. The answer is added to the Pareto set; therefore, the current neighborhood structure remained. Otherwise, go to the next neighborhoods' structure. It should be mentioned that once the newly generated answer is entered into the Pareto set, the algorithm starts with the first neighborhood structure again. In this situation, the algorithm replicates the search process for m times.

4) Evaluating the New Generated Answer.

In the evaluation stage of the new generated answer, one of the following states should occur:

First state

If f_1, f_2 for the new answers, in all answers of the dominant set, are worse, or one of them is worse and another is equal, it is not allowed to enter the dominant answer set.

Second state

The new answer is better than k answers of the dominant answer set. So, the new answer is added to the dominant set, and k nondominant answers will be eliminated.

Third state

The new answer is not better than any answers of the dominant answer set. Meanwhile, if one of the functions f_1 or f_2 is superior to all answers of the dominant answer set, it will enter into the Pareto set. However, other answers will remain in the dominant answer as well.

5) Replicating Algorithm

Steps 1 to 3 are replicated for m times.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] A. Müller, M. Weigl, B. Heiden, B. Herbig, J. Glaser, and P. Angerer, "Selection, optimization, and compensation in nursing: exploration of job-specific strategies, scale development, and age-specific associations to work ability," *Journal of Advanced Nursing*, vol. 69, no. 7, pp. 1630–1642, 2013.
- [2] V. G. Kulkarni, *Modeling and Analysis of Stochastic Systems*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2016.
- [3] A. A. Rajaei, A. Amindoust, and M. Asadpour, "Applying simulated annealing algorithm for parallel machine tardiness problem subject to job splitting," in *Proceedings of the 48th International Conference on Computers & Industrial Engineering (CIE48)*, pp. 2–5, Auckland, New Zealand, December 2018.
- [4] J. Puente, A. Gómez, I. Fernández, and P. Priore, "Medical doctor rostering problem in a hospital emergency department by means of genetic algorithms," *Computers & Industrial Engineering*, vol. 56, no. 4, pp. 1232–1242, 2009.
- [5] P. Rerkjirattikal, V. N. Huynh, S. Olapiriyakul, and T. Supnithi, "A goal programming approach to nurse scheduling with individual preference satisfaction," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2379091, 11 pages, 2020.
- [6] J. Bridges, P. Griffiths, E. Oliver, and R. M. Pickering, "Hospital nurse staffing and staff-patient interactions: an observational study," *BMJ Quality & Safety*, vol. 28, no. 9, pp. 706–713, 2019.
- [7] L.-F. Liu, S. Lee, P.-F. Chia, S.-C. Chi, and Y.-C. Yin, "Exploring the association between nurse workload and nurse-sensitive patient safety outcome indicators," *Journal of Nursing Research*, vol. 20, no. 4, pp. 300–309, 2012.
- [8] J. A. Baker, K. Canvin, and K. Berzins, "The relationship between workforce characteristics and perception of quality of care in mental health: a qualitative study," *International Journal of Nursing Studies*, vol. 100, Article ID 103412, 2019.
- [9] M. Hamid, R. Tavakkoli-Moghaddam, F. Golpaygani, and B. Vahedi-Nouri, "A multi-objective model for a nurse scheduling problem by emphasizing human factors," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 234, no. 2, pp. 179–199, 2020.
- [10] C.-C. Tsai and S. H. A. Li, "A two-stage modeling with genetic algorithms for the nurse scheduling problem," *Expert Systems with Applications*, vol. 36, no. 5, pp. 9506–9512, 2009.
- [11] S. Topaloglu and H. Selim, "Nurse scheduling using fuzzy modeling approach," *Fuzzy Sets and Systems*, vol. 161, no. 11, pp. 1543–1563, 2010.
- [12] S. M. Abdollahi and Z. Ansari, "A data-driven goal programming model for the nurse scheduling problem," *International Journal of Experimental Design and Process Optimisation*, vol. 3, no. 3, pp. 294–310, 2013.
- [13] T. C. Wong, M. Xu, and K. S. Chin, "A two-stage heuristic approach for nurse scheduling problem: a case study in an emergency department," *Computers & Operations Research*, vol. 51, pp. 99–110, 2014.
- [14] H. Jafari, S. Bateni, P. Daneshvar, S. Bateni, and H. Mahdioun, "Fuzzy mathematical modeling approach for the nurse scheduling problem: a case study," *International Journal of Fuzzy Systems*, vol. 18, no. 2, pp. 320–332, 2016.
- [15] A. Legrain, J. Omer, and S. Rosat, "An online stochastic algorithm for a dynamic nurse scheduling problem," *European Journal of Operational Research*, vol. 285, no. 1, pp. 196–210, 2020.
- [16] M. Vanhoucke and B. Maenhout, "On the characterization and generation of nurse scheduling problem instances," *European Journal of Operational Research*, vol. 196, no. 2, pp. 457–467, 2009.
- [17] Z. Liu, Z. Liu, Z. Zhu, Y. Shen, and J. Dong, "Simulated annealing for a multi-level nurse rostering problem in hemodialysis service," *Applied Soft Computing*, vol. 64, pp. 148–160, 2018.
- [18] A. Farasat and A. G. Nikolaev, "Signed social structure optimization for shift assignment in the nurse scheduling problem," *Socio-Economic Planning Sciences*, vol. 56, pp. 3–13, 2016.
- [19] S. Zanda, P. Zuddas, and C. Seatzu, "Long term nurse scheduling via a decision support system based on linear integer programming: a case study at the University Hospital in Cagliari," *Computers & Industrial Engineering*, vol. 126, pp. 337–347, 2018.
- [20] P. K. Nahand, M. Hamid, M. Bastan, and A. Mollajan, "Human resource management: new approach to nurse scheduling by considering human error," *International Journal of System Assurance Engineering and Management*, vol. 10, no. 6, pp. 1429–1443, 2019.

- [21] L. Wolbeck, N. Klierer, and I. Marques, "Fair shift change penalization scheme for nurse rescheduling problems," *European Journal of Operational Research*, pp. 1121–1135, 2020.
- [22] G. J. Lim, A. Mobasher, J. F. Bard, and A. Najjarbashi, "Nurse scheduling with lunch break assignments in operating suites," *Operations Research for Health Care*, vol. 10, pp. 35–48, 2016.
- [23] C. Di Martinelly and N. Meskens, "A bi-objective integrated approach to building surgical teams and nurse schedule rosters to maximise surgical team affinities and minimise nurses' idle time," *International Journal of Production Economics*, vol. 191, pp. 323–334, 2017.
- [24] A. J. Chiang, A. Jeang, P. C. Chiang, P. S. Chiang, and C. P. Chung, "Multi-objective optimization for simultaneous operating room and nursing unit scheduling," *International Journal of Engineering Business Management*, vol. 11, Article ID 1847979019891022, 2019.
- [25] M. Heshmat, K. Nakata, and A. Eltawil, "Solving the patient appointment scheduling problem in outpatient chemotherapy clinics using clustering and mathematical programming," *Computers & Industrial Engineering*, vol. 124, pp. 347–358, 2018.
- [26] P. Sarkar, D. Sinha, and R. Chaki, "A framework for solution to nurse assignment problem in health care with variable demand," in *Advanced Computing and Systems for Security*, pp. 3–20, Springer, Singapore, 2018.
- [27] A. Fügner, A. Pahr, and J. O. Brunner, "Mid-term nurse rostering considering cross-training effects," *International Journal of Production Economics*, vol. 196, pp. 176–187, 2018.
- [28] F. He, T. Chausalet, and R. Qu, "Controlling understaffing with conditional Value-at-Risk constraint for an integrated nurse scheduling problem under patient demand uncertainty," *Operations Research Perspectives*, vol. 6, Article ID 100119, 2019.
- [29] J. Schoenfelder, K. M. Bretthauer, P. D. Wright, and E. Coe, "Nurse scheduling with quick-response methods: improving hospital performance, nurse workload, and patient experience," *European Journal of Operational Research*, vol. 283, no. 1, pp. 390–403, 2020.
- [30] M. A. Awadallah, A. L. A. Bolaji, and M. A. Al-Betar, "A hybrid artificial bee colony for a nurse rostering problem," *Applied Soft Computing*, vol. 35, pp. 726–739, 2015.
- [31] T.-H. Wu, J.-Y. Yeh, and Y.-M. Lee, "A particle swarm optimization approach with refinement procedure for nurse rostering problem," *Computers & Operations Research*, vol. 54, pp. 52–63, 2015.
- [32] M. Bagheri, A. Gholinejad Devin, and A. Izanloo, "An application of stochastic programming method for nurse scheduling problem in real word hospital," *Computers & Industrial Engineering*, vol. 96, pp. 192–200, 2016.
- [33] Z. Zheng, X. Liu, and X. Gong, "A simple randomized variable neighbourhood search for nurse rostering," *Computers & Industrial Engineering*, vol. 110, pp. 165–174, 2017.
- [34] E. Rahimian, K. Akartunali, and J. Levine, "A hybrid integer programming and variable neighbourhood search algorithm to solve nurse rostering problems," *European Journal of Operational Research*, vol. 258, no. 2, pp. 411–423, 2017.
- [35] M. A. Awadallah, M. A. Al-Betar, A. T. Khader, A. L. A. Bolaji, and M. Alkoffash, "Hybridization of harmony search with hill climbing for highly constrained nurse rostering problem," *Neural Computing and Applications*, vol. 28, no. 3, pp. 463–482, 2017.
- [36] F. Knust and L. Xie, "Simulated annealing approach to nurse rostering benchmark and real-world instances," *Annals of Operations Research*, vol. 272, no. 1-2, pp. 187–216, 2019.
- [37] A. M. Turhan and B. Bilgen, "A hybrid fix-and-optimize and simulated annealing approaches for nurse rostering problem," *Computers and Industrial Engineering*, vol. 145, Article ID 106531, 2020.
- [38] P. Strandmark, Y. Qu, and T. Curtois, "First-order linear programming in a column generation based heuristic approach to the nurse rostering problem," *Computers and Operations Research*, vol. 120, Article ID 104945, 2020.
- [39] M. Asadpour, O. Boyer, and R. Tavakkoli-Moghaddam, "A blood supply chain network with backup facilities considering blood groups and expiration date: a real-world application," *International Journal of Engineering*, vol. 34, 2021.
- [40] M. Asadpour, O. Boyer, and R. Tavakkoli-Moghaddam, "Designing blood supply chain network in disaster situation considering blood groups and expiration date," in *Proceedings of the 1st International Conference on Challenges and New Solutions in Industrial Engineering and Management and Accounting*, Sari, Iran, July 2020.
- [41] I. Bahrami, R. M. Ahari, and M. Asadpour, "A maximal covering facility location model for emergency services within M (t)/M/m/m queuing system," *Journal of Modelling in Management*, 2020, ahead-of-print.
- [42] C. Ricciardi, A. M. Ponsiglione, G. Converso, I. Santalucia, M. Triassi, and G. Improta, "Implementation and validation of a new method to model voluntary departures from emergency departments," *Authorea*, 2020, Preprints.
- [43] G. Improta, G. Guizzi, C. Ricciardi et al., "Agile six sigma in healthcare: case study at Santobono pediatric hospital," *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, p. 1052, 2020.
- [44] P. Hooshangi-Tabrizi, I. Contreras, N. Bhuiyan, and G. Batist, "Improving patient-care services at an oncology clinic using a flexible and adaptive scheduling procedure," *Expert Systems with Applications*, vol. 150, Article ID 113267, 2020.
- [45] K. Rahimpour, H. Shirouyehzad, M. Asadpour, and M. Karbasian, "A PCA-DEA method for organizational performance evaluation based on intellectual capital and employee loyalty," *Journal of Modelling in Management*, vol. 15, no. 4, pp. 1479–1513, 2020.
- [46] M. Asadpour and H. Shirouyehzad, "Performance evaluation and ranking of academy award winners for best original score applying data envelopment analysis: 1990–2016," *Operations Research Letters*, vol. 47, no. 5, pp. 371–376, 2019.
- [47] I. Ajripour, M. Asadpour, and L. Tabatabaie, "A model for organization performance management applying MCDM and BSC: a case study," *Journal of Applied Research on Industrial Engineering*, vol. 6, no. 1, pp. 52–70, 2019.
- [48] R. Jafarpisheh, M. Karbasian, and M. Asadpour, "A hybrid reliability-centered maintenance approach for mining transportation machines: a real case in Esfahan," *International Journal of Quality & Reliability Management*, 2020.
- [49] S. Mirjalili, "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks*, pp. 43–55, Springer, Cham, Switzerland, 2019.
- [50] X. S. Yang, *Nature-inspired Optimization Algorithms. Chapter 5 - Genetic Algorithms*, pp. 77–87, Elsevier, Amsterdam, Netherlands, 2014.

Review Article

Involvement of Machine Learning Tools in Healthcare Decision Making

**Senerath Mudalige Don Alexis Chinthaka Jayatilake  and
Gamage Upeksha Ganegoda **

Faculty of Information Technology, University of Moratuwa, Katubedda, Moratuwa, Sri Lanka

Correspondence should be addressed to Gamage Upeksha Ganegoda; upekshag@uom.lk

Received 30 October 2020; Revised 18 December 2020; Accepted 9 January 2021; Published 27 January 2021

Academic Editor: Massimo Martorelli

Copyright © 2021 Senerath Mudalige Don Alexis Chinthaka Jayatilake and Gamage Upeksha Ganegoda. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the present day, there are many diseases which need to be identified at their early stages to start relevant treatments. If not, they could be incurable and deadly. Due to this reason, there is a need of analysing complex medical data, medical reports, and medical images at a lesser time but with greater accuracy. There are even some instances where certain abnormalities cannot be directly recognized by humans. In healthcare for computational decision making, machine learning approaches are being used in these types of situations where a crucial data analysis needs to be performed on medical data to reveal hidden relationships or abnormalities which are not visible to humans. Implementing algorithms to perform such tasks itself is difficult, but what makes it even more challenging is to increase the accuracy of the algorithm while decreasing the required time for the algorithm to execute. In the early days, processing of large amount of medical data was an important task which resulted in machine learning being adapted in the biological domain. Since this happened, the biology and biomedical fields have been reaching higher levels by exploring more knowledge and identifying relationships which were never observed before. Reaching to its peak now the concern is being diverted towards treating patients not only based on the type of disease but also their genetics, which is known as precision medicine. Modifications in machine learning algorithms are being performed and tested daily to improve the performance of the algorithms in analysing and presenting more accurate information. In the healthcare field, starting from information extraction from medical documents until the prediction or diagnosis of a disease, machine learning has been involved. Medical imaging is a section that was greatly improved with the integration of machine learning algorithms to the field of computational biology. Nowadays, many disease diagnoses are being performed by medical image processing using machine learning algorithms. In addition, patient care, resource allocation, and research on treatments for various diseases are also being performed using machine learning-based computational decision making. Throughout this paper, various machine learning algorithms and approaches that are being used for decision making in the healthcare sector will be discussed along with the involvement of machine learning in healthcare applications in the current context. With the explored knowledge, it was evident that neural network-based deep learning methods have performed extremely well in the field of computational biology with the support of the high processing power of modern sophisticated computers and are being extensively applied because of their high predicting accuracy and reliability. When giving concern towards the big picture by combining the observations, it is noticeable that computational biology and biomedicine-based decision making in healthcare have now become dependent on machine learning algorithms, and thus they cannot be separated from the field of artificial intelligence.

1. Introduction

Artificial Intelligence includes approaches and techniques like machine learning, machine reasoning, and robotics. In this review, the main concern will be given towards machine

learning as it is the approach that is being applied using different techniques and algorithms in various healthcare activities. The use of machine learning to solve clinical problems is called revolutionary clinical decision making. When machine learning is being used in clinical decision

making, it implies that the system will perceive a given individual by collecting and interpreting data relevant for the health of that individual and it will reason on the data to suggest the best actions that need to be performed to maintain or improve the individual's health. In machine learning, the system needs to learn the context of the problem and the quality of data provided. Normally, these algorithms are not very strong and concrete at the beginning, but by performing a single and repetitive task, the algorithm becomes strong with increased number of previous experiences.

There are two distinct ways that a clinical decision-making solution interprets by comparing previous knowledge that is included in the dataset. One is the fast or intuitive approach which uses underlying clinical pattern recognition and is typically used in medical emergencies. But these have a higher possibility of being erroneous and provide incomplete perception. The other approach is the slow or reasoned method. It is deductive, deliberate, and needs greater intellectual, time, and cost information. But the decisions made are more accurate. As all these decisions are based on the data that are being collected, analysed, and stored in complex and heterogeneous forms, it is important to use algorithmic approaches to minimize the computing power required. Machine learning applications have currently contributed massively to the healthcare sector worldwide to improve its quality and will continue to do so [1].

It is also important to mention that when discussing computational decision making in the healthcare sector, it is not always about detecting or predicting diseases, biomedicine, biomedical image analysis, etc., but also about how to perform medical treatment research, patient care, allocating resources, managing hospital volume, public health policymaking, and much more. As an example, when considering the current situation that has arisen with COVID-19 pandemic, it could be seen that all the aforementioned points need to be considered in healthcare and those tasks need to be performed within a considerably small time period. Therefore, the best approach is to use machine learning-based decision making in the healthcare sector in such times. This is the reason why there is a demand for the area "emergency machine learning" in the present world [2].

Even though artificial intelligence is set to make transformation by playing an important role in healthcare, there are a few ethical aspects as well that need to be considered when implementing such systems and obtaining decisions from them. Few of the ethical issues are the accountability and transparency of the decisions made by such systems, the potential for group harms arising from algorithmic bias and the professional roles, and the integrity of the clinicians. Therefore, it is important to give consideration when implementing such systems and balance them with the benefit that they create with more efficient healthcare systems by the high and accurate computational power of artificial intelligence at a considerably low cost. Moreover, algorithms in artificial intelligence have the capacity to perform computerized predictive analysis by filtering, ordering, and searching for patterns from big datasets from

multiple sources to provide fast and informed decisions. In the present context, due to this discussed matter, most jurisdictions do not allow directly applying these algorithms to make the final decisions, but instead use them as an aid for diagnosis [3].

Throughout the paper many machine learning algorithms and their applications in the field of computational decision making for healthcare will be discussed along with the methods that are used to improve the efficiency of the algorithms with the objective to highlight the importance of scalable machine learning algorithms in healthcare applications. The aim of the paper is to discuss the involvement of machine learning algorithms in the healthcare sector to perform computational decision making starting from the initial phase where machine learning was introduced to computational biology till the peak it currently stands which is the introduction of precision medicine to the field of biomedicine. This paper is organized in different sections. Machine learning approaches and algorithms that are being applied in healthcare for decision making will be discussed in Section 2 followed by the applications of machine learning in the healthcare sector in various aspects such as disease prediction and detection, medical imaging, machine learning in biomedicine, biomedical event extraction, machine learning approaches to polypharmacology, and machine learning for drug repurposing using system biology. The discussion section will include an evaluation and comparison of machine learning algorithms with regard to the applications of these algorithms in healthcare which will be followed by recognition of different mechanisms used to enhance the accuracy of the algorithms within applications, along with further information on the involvement of scalable machine learning algorithms for computational decision making in healthcare. The latter part of the paper consists of the conclusion.

2. Machine Learning Approaches and Algorithms

Machine learning can be introduced as a scientific discipline that focuses on how computers learn from data and continuously improve themselves. It is mainly based on probability and statistics. But it is more powerful than the standard statistical methodologies when it comes to decision making. Information gathered from a dataset which is being given to the algorithm is called features. The accuracy of the predictions made by the model is dependent on the quality of the features provided to the algorithm. It is the duty of a machine learning developer to detect the subset of features that could best fit the purpose, increasing the accuracy of the model. This is not an easy task. Continuous experiments should be carried out to identify the said feature subset for the algorithm. When considering putting a machine learning algorithm to applications, there are basically three steps to follow, which are training, testing, and validation. Training is important as the accuracy of the results will be depending on the training dataset. Using the test dataset, the performance of the algorithm will be measured. When using the test data for measuring the performance, it is also

important to lower the bias and to increase the variance in this testing period. A good machine learning algorithm must optimize the bias-variance trade-off. The evaluation of the final machine learning algorithm performance is done based on the validation dataset in the validation period [4]. As a start, it would be better to have an idea about various approaches taken in machine learning along with several algorithms that are being used excessively for clustering and classification purposes in machine learning.

2.1. Supervised Learning. In supervised learning, a training set is provided with appropriate objectives in this approach. Classification and regression are the two categories found in supervised learning. In classification, with the use of classification methods, the trained system allocates inputs into classes. In regression, the sources are continuous rather than discrete. The root-mean-squared error is being used to evaluate regression predictions, while accuracy is being used to evaluate classification predictions. [5]. Supervised learning has the goal of predicting a known output based on a common dataset. Tasks performed by supervised learning can most of the time be performed by a trained person as well. Supervised learning focuses on classification which involves choosing among subgroups to best describe a new instance of data and prediction, which involves estimating an unknown parameter. This is often used to estimate and model risk while finding relationships which are not readily visible to humans [6]. Below are a few supervised learning algorithms which are widely used in the field of computational biology and biomedicine.

2.1.1. K-Nearest Neighbour (KNN). KNN is a popular supervised classification algorithm which is used in many fields such as pattern recognition, intrusion detection, and so on. KNN is a simple algorithm which is easy to understand. Even the accuracy is high in KNN, but the issues are that it is computationally expensive and it has a high memory requirement as both testing and training data need to be stored [7]. A prediction for a new instance is obtained by finding the most similar instances at first and then summarizing the output variable according to those similar instances. For regression, this can be the mean value, and for classification, this may be the mode value. To determine the similar instance, the distance measure is used. Euclidean distance is the most popular approach used to calculate the distance. The training dataset should be vectors in a multidimensional feature space, each with a class label [5].

2.1.2. Support Vector Machine (SVM). SVM is a supervised machine learning algorithm which is used to address mainly classification problems but also used for regression issues. In this algorithm, initially, the data items are plotted as points in an n-dimensional space with the feature value being the particular coordinate. Then, it identifies the hyperplane that separates the datapoints into two classes. By this, the marginal distance between the decision hyperplane and instances that are close to the boundary can be maximized

[5]. What brings SVM ahead of other algorithms is that it has basic functions that can map points to other dimensions by using nonlinear relationships [8]. As it divides the datapoints to two classes, SVM is also known as the non-probabilistic binary classifier. SVM has more accuracy when compared with many other algorithms. But it is best suited for problems with small datasets. The reason is that when the dataset keeps on getting larger, the training becomes more complex and time consuming. When data have noise, it cannot perform well. To make the classification more efficient, SVM uses a subset of training points. SVM is capable of solving both linear and nonlinear problems, but nonlinear SVM is preferred over linear SVM as it has better performance [7].

2.1.3. Decision Trees (DTs). DT is a supervised algorithm which has a tree like model where decisions, possible consequences, and their outcomes are being considered. Each node carries a question, and each branch represents an outcome. The leaf nodes are class labels. When a leaf node is being reached by a sample data, the label of the corresponding node will be assigned to the sample. This approach is suited when the problem is simple and when the dataset is small. Even though the algorithm is easy to understand, it has certain issues such as the overfitting problem and biased outcomes when working with imbalanced datasets. But DT is capable of mapping both linear and nonlinear relationships [7].

2.1.4. Classification and Regression Trees (CARTs). CART is a predictive model from which the output value is predicted based on the existing values in the constructed tree. The representation for the CART model is a binary tree in which each root represents a single input and a split point on that variable. Leaf nodes contain an output which is used to make predictions [5].

2.1.5. Logistic Regression (LR). LR is a popular mathematical modelling procedure which is used for epidemiologic datasets in the area of machine learning. It first calculates using the logistic function. Then, it learns the coefficients for the logistic regression model and then finally makes predictions using that logistic regression model [9]. This model is a generalized linear model and has two parts, namely, linear part and link function. The linear part is responsible for carrying out the calculations of the classification model, and the link function is responsible for delivering the output of the calculation [10]. LR is a supervised machine learning algorithm which needs a hypothesis and a cost function. It is to be noted that optimizing the cost function is important [11].

2.1.6. Random Forest Algorithm (RFA). RFA is a trending machine learning technique which is capable of both regression and classification [12]. It is a supervised learning algorithm in which the ground methodology is recursion. In this algorithm, a group of decision trees are being created

and the bagging method is used for training purposes [13]. RFA is insensitive to noise and can be used for imbalanced datasets. The problem of overfitting is also not prominent in RFA [7].

2.1.7. Naive Bayes (NB). NB is a classification algorithm which is used for binary and multiclass problems. The NB classifiers are a collection of classifying algorithms that are based on the Bayes theorem. But they all adhere to a common principle which is every pair of features being classified must be independent of each other [5]. This is a bit similar to SVM, but the process takes advantage from statistical methods. In this method, when there is a new input, the probabilistic value will be calculated among the classes with regard to the given input and the data will be labelled with the class which has the highest probabilistic value for the given input [9].

2.1.8. Artificial Neural Network (ANN). ANN is a supervised machine learning approach which is well known for image classification problems. In machine learning, artificial neurons are considered to be the basic concept of ANN and it is similar to a biological neural network. There are 3 layers in an ANN, and every node in each layer is connected with all the nodes in the other layers. By increasing the number of hidden layers, a deeper neural network can be created [14]. In neural networks, there are three types of functions. Error function will determine how good or bad the output was for a given set of inputs. The search function will identify the changes that would reduce the error function. Update function will determine how the changes will be made as per the search function. This is an iterative process that would improve the performance of the algorithm [8].

2.2. Unsupervised Learning. When a developer does not have a clear understanding of the data that are involved with the system, it is not possible to label the data and provide them as the training dataset. In these cases, the machine learning algorithms themselves can be used to detect similarities and differences between the data objects. This is the unsupervised approach of machine learning. In this method, existing patterns will be identified and the data will be clustered according to the identified patterns [4]. Therefore, in unsupervised learning, the system makes decisions without being trained by a dataset as no labelled data are being given to the system which could be used for predictions [5]. It is to be noted that unsupervised learning is an attempt to find naturally occurring patterns or groups within data. The challenging part in it is to find whether the recognized patterns or groups are useful in some way. This is the reason for unsupervised learning to play a major role in precision medicine. As a simple example, when grouping individuals according to their genetics, environment, and medical history, certain relationships among them which were not visible before might get identified by unsupervised machine learning algorithms [6]. K-means, mean shift, affinity propagation, density-based spatial clustering of applications

with noise (DBSCAN), Gaussian mixture modelling, Markov random fields, iterative self-organizing data (ISO-DATA), and fuzzy C-means systems are a few examples for unsupervised algorithms [8].

Clustering is an approach in unsupervised learning, and it can be used for dividing inputs into clusters. But these clusters are not identified initially but are grouped based on resemblance [5]. In clustering, the root approaches are separated as per the different features that they carry. They can be partitioning (k-means), hierarchical, grid-based, density-based, or model-based, and they can be further divided as numerical, discrete, and mixed data types. Inheritance relationships between clustering algorithms within an approach show common features and improvements that they make on each other. Speed, minimal parameters, robustness to noise, outliers, redundancy handling, and object order independence are the desired clustering features which are required in a clustering algorithm to be implemented within a biomedical application [15]. Clustering algorithms are used when datasets are too large and complex for manual analysis. Therefore, they must be fast and they must not be affected by redundant sequences. It is important to evaluate a clustering algorithm to know whether it is suitable for the problem in hand by considering features such as

- (1) Scalability: runtime and memory requirements should not exceed when working with large datasets.
- (2) Robustness: ability to detect outliers that are different from the rest of the samples.
- (3) Order insensitivity: the order that the inputs are given should not be affecting the final output.
- (4) Minimum user-specified input: the minimum number of parameters that should be provided.
- (5) Mixed data types: the objects provided may have different data types.
- (6) Arbitrary-shaped cluster: ability to find arbitrary-shaped clusters.
- (7) Point proportion admissibility: duplicating objects and reclustering should not change the result [15].

When giving concern to the above factors, it is to be noted that as a result of the use of unsupervised machine learning in healthcare, when a patient is being diagnosed for a specific disease, in that process itself the patient will be identified of another disease as well if any such disease is present. The reason is that the algorithm has learnt itself on facts that need to be considered for various types of diseases and it analyses the given data and categorizes as diseases according to those data. In addition, even though various reports taken to identify diseases may include different types of data, using this approach, all these data can be analysed simultaneously which is of much convenience and also time saving. With regard to the aforementioned aspects, there is no doubt that healthcare decision making is highly benefited from the unsupervised machine learning approach.

2.2.1. Partition Clustering. In partition clustering, the objects are partitioned and may change clusters based on the

dissimilarity. It is useful in bioinformatics when the number of clusters is decided such as for a small gene expression dataset. The drawback is that the user needs to manually enter the number of clusters as an input. But this approach is commonly used in bioinformatics. Fuzzy k-means, COOLCAT, clustering large applications (CLARA), and clustering large applications based on randomized search (CLARANS) are a few examples for partition clustering algorithms [15].

2.2.2. Graph-Based Clustering. Graph-based clustering is used in interactomes to make complex predictions and to sequence networks. This approach is often slow and sensitive to user-specified parameters. super-paramagnetic clustering (SPC), Markov cluster algorithm (MCL), molecular complex detection (MCODE), and restricted neighbourhood search cluster (RNSC) are a few examples for graph-based clustering algorithms [15].

2.2.3. Hierarchical Clustering. In hierarchical clustering, the objects are partitioned into a tree of nodes and these nodes are considered as clusters. There are parent nodes and child nodes. A node could have just one parent, and each node can have zero or more child nodes. This approach is popular in bioinformatics as clusters can be navigated at various levels of granularity. The drawbacks are that they are often slow, errors made when merging clusters cannot be undone even though it affects the result, and if large clusters are merged, then interesting local cluster structure may be lost. This approach is used to represent protein sequence family relationships and also could be used to show gene relations reflecting their gene similarity. Chameleon, robust clustering using links (ROCK), scalable information bottleneck (LIMBO), and spectral are a few examples for hierarchical clustering algorithms [15].

2.2.4. Density-Based Clustering. Density-based clustering uses a local density criterion, and the clusters are subspaces in which the objects are dense and are separated by subspaces of low density. It is used in bioinformatics to find the densest subspaces in interactome networks, typically involving cliques. Time efficiency and the ability to find clusters of arbitrary shapes are the advantages of this approach. Some of these algorithms accept user parameters, but it is not the number of clusters. Ordering points to identify the clustering structure (OPTICS), clustering in quest (CLIQUE), density based clustering (DENCLUE), and clustering categorical data using summaries (CACTUS) are a few examples for density-based clustering algorithms [15].

2.2.5. Model-Based Clustering. In model-based clustering, it is assumed that objects match a model which is often a statistical distribution. The model can be user specified using a parameter, and this model can even be changed in the process. This approach can be found in bioinformatics to integrate background knowledge into gene expressions, interactomes, and sequences. Slow processing time on large

datasets is a drawback of this method. If the user assumptions are false when defining the models, then the results will also be inaccurate. SVM-based clustering, COBWEB, and AutoClass are a few model-based clustering algorithms [15].

2.3. Semisupervised Learning. For semisupervised learning, a partial training set of data is provided. This type of training is used when some missing results could be targeted by some training data. Semisupervised learning algorithms are trained on both labelled and unlabelled data. Due to this reason, it exhibits the features of both supervised and unsupervised machine learning algorithms [16].

2.4. Evolutionary Learning. Evolutionary learning is mainly used in the biology field to learn about biological organisms and predict their survival rate. Using this method, the level of correctness of a result can also be predicted [16].

2.5. Active Learning. In active learning, the system gets the training tags only for a restricted set of occurrences. By using it, the optimality of the substances can be enhanced to gain tags for the required goal. The advantage in this approach is that the algorithm not only continuously learns but also gets the facts which were self-learned approved either by querying a user or an information source in an interactive manner. It is something similar to budget functions in an organization and is a modern machine learning approach for decision making [16].

2.6. Deep Learning. Deep learning is an advanced phase of machine learning which evolves around neural networks for learning and predicting data. Using this approach, complex generalized systems can be implemented which are able to accept any type of problem and give predictions regarding it [16].

2.7. Reinforcement Learning. In reinforcement learning, the training data are provided only as a response to the program's activities in a self-motivated situation. It has a continuous learning process from the environment in an iterative fashion [13].

After discussing several machine learning approaches, it would be better to list down a few examples on the applications of machine learning in the field of biomedicine so that this review will be interesting from the very beginning. In neuroscience, machine learning classifiers are being used to study functional and structural dynamics of the brain. Machine learning approaches are used in cancer prediction and prognosis. SVM classifiers are used to detect prostate cancer. Hierarchical clustering has been used in investigation of Alzheimer's disease. ANN has been used in classifying different subtypes of psychogenic nonepileptic seizures [4]. With the knowledge gathered on various machine learning approaches and machine learning algorithms which are mostly connected with computational biology and biomedicine, now it is time to dive into deeper knowledge

and to identify the applications of these algorithms in the discussing field.

3. Machine Learning in Disease Prediction and Detection

Various machine learning approaches have been implemented to predict or detect a disease at its early stages so that the treatment for it would be less complex and it would increase the probability of the patient being cured. As a result of these approaches, different types of diseases have been detected but with diverse accuracy levels depending on factors such as the used algorithm, feature set, training dataset, and so on. In this section, a few selected diseases will be discussed as examples, along with the importance of identifying a disease at the earliest, the machine learning methods implemented to detect the disease, and the features that were considered to make predictions. A descriptive comparison of the machine learning approaches which have been implemented will be conducted in the discussion section of the paper, followed by suggestions to further improve them.

3.1. Cancer. Human body has the right count of cells of each type. Cancer begins with abrupt changes in the cell organization. Signals which are being generated by cells determine the control and division of cells. When these signals become faulty, cells multiply too much which form a lump called tumour. Nowadays, thermography is more reliable as it is noninvasive and nonionizing. With the emerging technology, it has been producing efficient and positive results which have made it superior over other technologies. From the thermographic images, with the use of feature extraction techniques and machine learning techniques, the presence of cancer cells can be detected. Scale invariant feature transform (SIFT) and speeded up robust feature (SURF) techniques can be used to extract features from images. Using principal component analysis (PCA), the features could be further filtered in order to make better interpretations [7].

3.1.1. Breast Cancer. Breast cancer is a type of cancer that is mostly seen in women and is a leading cause for women's death. But this can be reduced by early detection of cancerous cells by tests like magnetic resonance imaging (MRI), mammogram, ultrasound, and biopsy. Breast cancer is diagnosed by classifying the tumour. Tumours can be either benign or malignant. It is to be noted that malignant tumours are more harmful than benign tumours. But it is not an easy task for physicians to distinguish among these tumours. This makes machine learning algorithms important as they can automatically learn and improve from the experiences without being explicitly programmed [5].

In the past years, many machine learning techniques were developed for breast cancer detection and classification. Their process could be analysed in three stages which are preprocessing, feature extraction, and classification. Feature extraction stage is important as it helps in

discriminating between benign and malignant tumours. Then, the image properties such as smoothness, coarseness, depth, and regularity are extracted using segmentation [17].

Normally, images are converted to binary to extract useful information. But it has been observed that once doing so, some important features in the image vanished which omits crucial information. This has led to keeping the images in the grey scale format. Using discrete wavelet transformation (DWT), the images can be transformed from the time domain to the frequency domain. This wavelet decomposition contains four matrices which are the approximation coefficient matrix, the horizontal detailed coefficient matrix, the vertical detailed coefficient matrix, and the diagonal detailed coefficient matrix. These are the values that will be used for the machine learning algorithms [11].

3.1.2. Lung Cancer. Lung cancer can initiate in the windpipe, main airway, or lungs. People with emphysema and previous chest problems have a higher probability of being diagnosed with lung cancer. Tobacco, smoking, and air pollution can be a few major risk factors for lung cancer. Lung cancer starts in the lungs at the primary stage and spreads to other organs as the secondary stage. Symptoms of lung cancer will not be shown until the disease is quite advanced. That is what makes it more dangerous [9].

Computerized tomography (CT) reports are less noisy as compared to MRI and X-ray reports. Grayscale conversion, noise reduction, binarization, and segmentation techniques are important to get the image in the required form with less noise and distortion. When converting to grey scale, the average of RGB is taken. The median filter is used for noise reduction. Segmentation removes unnecessary details from the images and locates the objects and the boundaries. In feature extraction stage, features such as area, perimeter, and eccentricity are considered [18].

Small-cell lung cancer (SCLC) detection is extremely difficult for human as it is almost identical to the one without. This is where the machine learning algorithms such as convolution neural network- (CNN-) based deep learning methods could be used in detecting SCLC. Usually, deep learning algorithms require large training datasets which is an issue. Entropy degradation method (EDM) can be used to overcome the said matter. The training data and testing data need to be high-resolution lung CT scans. EDM carries the concept of shallow neural network where vectorized histograms are converted to scores. Then, the scores are transformed to probability using logistic function. In this approach, SCLC detection is considered as a binomial problem which contains only two groups: either a healthy person or a lung cancer patient. So, initially test data are also given with both these types. This approach is reasonably accurate but not the best, and there is a large space to be further improved. But it is recommended that it could be further improved by providing a larger training set and a deeper network. By combining with CNN, the image processing is also further improved for better detection as CNN is being used in many applications of CT imaging [19].

3.1.3. Acute Lymphoblastic Leukaemia. Acute lymphoblastic leukaemia (ALL) is a type of cancer where a large number of immature lymphocyte blood cells develop and they affect the production of other blood cells. This progresses rapidly and can be very fatal within a month or a week. Pale colour of skin, patient feeling very tired, lymph node getting enlarged, fever, and joint pain are a few symptoms that were identified in the patients who were diagnosed with ALL. Machine learning algorithms play a vital role when trying to automatically segment and classify microscopic images to detect leukaemia.

There have been various machine learning algorithms used for leukaemia detection such as KNN, SVM, NB, radial basis function network (RBFN), and multilayer perceptron (MLP). But in all these approaches, there are basically four sections which are preprocessing, feature extraction, classification model building, and evaluation of the classifier. In the preprocessing stage, cropping of the image will be done so that the region of interest (ROI) is clearly visible and the unwanted information is eliminated. Using the Gaussian blur smoothing technique, the images can be further processed to enhance the picture by reducing the noise. In the feature extraction stage, concern is given towards colour-based features, geometrical features, statistical features, Haralick texture feature, image moments, local binary pattern, and presence of adjacent cells [20].

3.2. Diabetes. Diabetes is a chronic disease, and it needs to be identified at the early stages for correct medication. Diabetes is caused when the sugar ratio in blood increases. This makes the life complicated for the patients due to many reasons. Diabetes can be classified under three types, namely, diabetes 1, diabetes 2, and gestation diabetes.

Discriminant analysis (DA) is a procedure in which the class label of an input is determined by a series of equations that are obtained by input features. Normally, DA uses two possible objectives which are finding a related equation for classifying test samples and interpretation of the predictive equation to better understand the relationship among features. Being pregnant, the weight of the patient, blood pressure, glucose concentration, the ratio of insulin in blood, diabetes pedigree function (DPF), skinfold thickness, and patient age are some features which can be considered for the classification [10].

By using machine learning algorithms such as Gaussian Naive Bayes (GNB), LR, KNN, CART, RFA, and SVM along with variables in electronic medical records (EMRs) such as serum-glucose1, serum-glucose2 levels, body mass index (BMI), age, race, gender, creatinine level, and so on, prediction of type 2 diabetics was possible [21]. Time to time various machine learning techniques were used to try and improve the accuracy of the predictions made. One approach was made using neural networks. In this method, a feed-forward neural network was trained by a back-propagation algorithm. Even in this approach, the features that were considered are the number of pregnancies, skinfold thickness, serum-insulin, BMI, DPF, and age and the main feature considered was the plasma-glucose level. It

was observed that the predictions made by using neural networks showed a higher accuracy, when compared with other machine learning algorithms [22]. Research has been carried out on using deep neural networks (DNNs) as well for prediction of diabetics by training the DNN using five-fold and ten-fold cross validation. It is to be highlighted that both aforementioned approaches which were taken using neural networks have shown an accuracy near 97% in diabetes prediction [23].

3.3. Heart Diseases. Heart diseases are severe events which are caused by blockage inside the heart arteries. Chronic heart disease is the rise of plaque inside the coronary arteries. This progresses slowly and could lead to a heart attack. Peculiar glucose metabolism, extreme blood pressure, dyslipidaemia, smoking, lack of physical exercise, and age are few risk factors that have been identified for major heart diseases. Symptoms of a heart disease may include shortness of breath, weakness of physical body, swollen feet and fatigue with related signs, and so on [14].

In the field of cardiology, the tasks that precision medicine has performed include diagnostics as well as therapeutics in various subfields. Interventional cardiology, personalized treatment options in correcting heart rhythms, some gender differences affecting the outcome of cardiovascular diseases, and numerous works done in genomics can be highlighted as areas in which tasks have been performed by precision medicine in cardiology. Nowadays, in healthcare informatics, there are services such as patient monitoring and clinical decision support systems (CDSSs). With the advancement of machine learning, now complex problems can be solved even by machines which was only possible by humans, decades back. By utilizing these techniques in precision medicine, the CDSS could be modified to make complex clinical decisions, recognizing newer phenotypes and planning person-oriented specialized treatment options.

In cardiology, blood tests are popular among the different investigation methods in precision medicine. AGES is a specific precision medicine test which utilizes other factors in addition to blood tests to avoid ischemic heart disease. In precision medicine, concern is mainly given towards genetics and there are many research studies being carried out to find genetic causes of a disease. Cardiac genetics, cardiac oncology, and ischemic heart disease can be identified as specific areas of interest in precision cardiology. Methods such as blood tests, genetics tests, image tests, or even a combination of them may be used for diagnostic and therapeutic purposes when advancing with precision medicine in cardiology. Many cardiovascular diseases have their roots embedded in genetics. Therefore, solutions using precision medicine are considered more productive specially in these types of diseases. CNN, recurrent neural network (RNN), natural language processing (NLP), SVM, and long short-term memory (LSTM) are few machine learning techniques which could be used efficiently to make precise CDSS using deep learning [24].

A machine learning approach to identify cardiologic diseases includes preprocessing, feature selection, cross validation method, machine learning classifiers, and classifier performance evaluation. There are several preprocessing techniques such as removing of missing values, standard scalar, and MinMax scalar. Feature selection is important in machine learning as irrelevant features can affect the classification performance of the algorithm. By applying feature selection prior to the classification, the execution time is reduced and the accuracy of the classification is increased. There are various feature selection algorithms. Relief, Minimal Redundancy Maximal Relevance (mRMR), and Least Absolute Shrinkage and Selection Operator (LASSO) are a few popular feature selection algorithms [25].

3.4. Chronic Kidney Disease (CKD). CKD is a type of a kidney disease which gradually affects the kidney functionality and leads to kidney failure. CDK can be diagnosed using clinical data, lab tests, imaging studies, and biopsy. But biopsy has some disadvantages such as being invasive, costly, time-consuming, and sometimes being risky. This is where machine learning can be applied to overcome the aforementioned disadvantages. In many disease predictions using machine learning, SVM was a commonly used classifier. But for CKD, there is not much research that could be found that uses SVM for the classification. ANN, DT, and LR were the main machine learning classifiers used in this domain. When observing the obtained results, ANN showed far better performance when compared with DT and LR on CKD diagnosis [26].

3.5. Parkinson's Disease (PD). PD is a chronic and progressive movement disorder. It has no causes, no permanent cure, and limited treatment options. It is found that PD occurs due to reduced production of dopamine, a chemical that controls movement and coordination. Tremors, rigidity, slowness of movement, and postural instability are a few symptoms of PD. Abnormal writhing movement is a significant symptom of this disease. Some researchers have applied machine learning algorithms on video recordings and computer vision to differentiate healthy controls from PD patients. Some researchers have also used voice samples to differentiate healthy controls from PD patients [27]. PD belongs to the neurodegenerative disease category that may directly or indirectly affect the brain cells which will result in affecting movement, speech, and other cognitive parts [28].

The feature set for the classification can be obtained based on PCA and genetic algorithm (GA). GA is inspired by Darwin's theory of evolution. A variable is considered as a gene. Chromosomes are a sequence of a gene. A predefined function evaluates the quality of a chromosome and the high-performing chromosome will be used to create the offspring. Genetic operations such as mutation and cross-over are used to create the offspring. Basically, it is a competition between the chromosomes in which the fittest will survive till the end. This is the concept which will be behind the feature extraction using the GA [29]. PCA is an

unsupervised linear conversion technique and a statistical technique commonly used to identify new patterns in high-dimensional data. Common applications that use PCA are face recognition and image compression [30].

3.6. Dermatological Diseases. Dermatological diseases are complex and have a large variety, and people have scarce expertise on it. Early detection is always preferred as it could lead to serious outcomes. Eczema, herpes, melanoma, and psoriasis are a few dermatological diseases which should be identified at the early stage to take life out of danger.

In one approach that was taken to diagnose dermatological diseases, the first phase involved data collection and data augmentation using images. Phase 2 is very important as it is where the model is created and trained. In the last phase, the image is converted to an array and the features are broken down using the trained model that was created. There are various augmentation techniques such as synthetic minority oversampling technique (SMOTE) and computer vision techniques like grey scaling, blurring, increasing contrast, changing the colour channel, sharpening, reducing the noise, and smoothing. When there are more instances in the database, it is better for the training of the model. Training the CNN with a large dataset overcomes the overfitting problem. SVM classifier is used for the prediction. The features in the final convolutional layer can be directly given to the SVM as an input. But in order to do this, the SVM must be trained by giving the trained features from the final convolutional layer as the training dataset. Then, the SVM will convert it to vectors and store them [31].

4. Machine Learning in Medical Imaging

Medical imaging is a rapidly growing research area as it is important to diagnose diseases in many instances. Several steps can be identified when observing the process of making predictions from an image using machine learning. Once an image is given as the input, it will be divided into different segments to zoom the interested area. Then, with the use of information retrieval techniques, features can be extracted from those areas. Among them, the required features are selected and the noise is removed. Finally, the classifier will classify the extracted data and will make predictions based on the classification.

Nowadays, in the medical community, accurate diagnosis of a disease by processing large amounts of medical data is crucial. In the field of medicine and biology, there are various tasks which are being carried out using machine learning algorithms. Distribution of data on the basis of their characteristics, medical data examination, disease diagnosis and treatment planning, data gathering and inspection, correcting diagnostic of different diseases by medical imaging, and extracting features from medical images on diseases are just a few of the applications.

When further consideration is given towards the applications of medical imaging, it can be observed that medical imaging is extensively being used to improve planning of surgical procedures with regard to many

diseases. Therefore, before discussing the involvement of machine learning in medical imaging, below mentioned are a few such applications that would better explain how medical imaging is being applied in surgical planning to obtain positive results while mitigating the risks.

As the skull is one of the most complex areas in a human body in both anatomical and surgical perspectives, surgical management is extremely difficult when working with a wide variety of lesions. But during the last decades, the endonasal endoscopic route was considered as a suitable approach for several skull-based lesions. A main advantage is that when performing a skull-based surgery through the nose with the aid of an endoscope, direct visualization of neurovascular structures of different areas of the skull base can be obtained with minimum brain displacement and manipulation. The endoscope also provides a wider and multiangle close-up view which is of much importance in the surgical field [32]. For diseases like rectal cancer, MRI plays a key role as it can accurately depict the local extent of the cancer and generates relevant information required for prognoses which can directly influence the choice of the optimal therapeutic procedure used for each individual patient which encourages the area of personalized medicine [33].

It is to be noted that when image-guided surgeries are compared with conventional surgical approaches, image-guided surgeries are less invasive, have more precise targeting, and have improved outcomes. Imaging is used to plan, monitor progress, and to assess results. In neurosurgery, this is more crucial as the object of the surgical procedure is hard to be located and needs to be reached with minimum damage to the healthy tissues. Apart from it, to guide the placement of the surgical instruments and to ensure that the required tissue is being treated, medical imaging is vital. At present, a wide variety of medical imaging modalities are being used such as magnetic resonance imaging, computed tomography, ultrasound, positron emission tomography, single-photon emission computed tomography, fluoroscopy, and so on to perform different assessments like biopsies, tumour resection, epilepsy, vascular conditions, and so on [34]. Medical image analysis is evolving day by day with the development of technology. This is further assisted by 3D virtual model creation as well to improve understanding of complex anatomy and to provide powerful tools for surgical planning and intra-operative guidance. Nowadays, the use of 3D ultrasound and fetal MRI is also becoming common in clinical practice [35]. The interesting fact is how machine learning could be integrated with these medical imaging techniques to make better decisions in surgical planning. It is to be noted that using machine learning, it is possible to understand hidden relationships which may not directly be visible to humans when observing multiple data. This is the reason that even disease predictions in healthcare have been made possible through machine learning approaches. Therefore, by processing these medical image data using unsupervised machine learning techniques such as clustering, an analysis could be performed on the dataset which could later be referred by the surgeon to identify if any crucial information has been lost while planning the surgery or else to even

confirm that the decisions made on the approach of conducting the surgery is suitable for the considered patient.

Entities such as lesions and organs in medical images are too complicated and they cannot be shown correctly with the use of simple mathematic solutions. The pixel analysis in machine learning is used for medical image processing from which certain values can be extracted directly from the image. Feature calculation and segmentation is not required in pixel-based machine learning. Due to this, even an image with low contrast will have no issue in processing to extract information. Pixel analysis utilizes longer training time because of the high dimensionality of data. Histogram equalization (HE) is an efficient technique which could be used for contrast improvement. There are various other extensions of HE which were implemented in order to improve the performance of the algorithm. Linear discriminant analysis (LDA), SVM, and DT are few machine learning methods used to analyse medical images. Low binary pattern descriptors are being implemented using machine learning approaches which could be used on biological images. The neural network technique is used in medical images to investigate the details regarding a disease. Machine learning in medical imaging can also be found in medical-based expert systems [16].

CNN is one of the best models for image analysis. It has several layers that could convert the input by using convolution filters. There are two sections in classification with regard to medical images. They are image classification and object classification. In image classification, deep learning is used to investigate clinical-related issues so that early treatment for the patient will be possible. In object classification, the main target is to analyse more on interested small chunks of the medical image. In medical image analysis, deep learning algorithms help in categorizing, classifying, and enumerating disease patterns using image processing [16].

Computers have the ability of performing tasks consistently and tirelessly. In the past years, machine learning has proved its ability to learn and master tasks which were considered as too complex for machines to handle, and in some instances, they have been able to identify patterns which are beyond human perception. When using machine learning on medical images, there are few terms which are widely used which need to be discussed. Classification means labelling or assigning a class to a group of pixels. Model is a set of decision points which are learnt by the machine learning algorithm. Algorithm is the steps which have been taken to implement the model. Labelled data are the set of data which are examples for a specific class. Validation set or the training set is the set of data which is used to train the system. Node is a part of a neural network which includes two or more inputs and an activation function. A layer is a combination of nodes from which the outputs are computed. Segmentation is a process of splitting the image into sections so that more focus could be given to the split segments. Overfitting is when a classifier is too specific to the training set and it is not useful as it is only familiar with those examples. Having many features can lead to overfitting. Features are the numeric values that represent the example.

When it is with regard to medical images, it can be actual pixel values, edge strengths, variation in pixel values in a region, etc. Feature selection must be done in a way that the selected subset of features is able to provide the best and the most accurate predictions [8].

Image recognition and biomedical time series classification are nonlinear classification problems. With the existing classification algorithms and feature extraction technologies, it is not possible to have highly complicated nonlinear functions. But using DNN, nonlinear functions can be constructed by increasing the number of layers and neurons in the network. With the use of ensemble learning, multiple classifiers could also be combined to have complicated decision-making functions. Both in SVM and ensemble learning, nonlinear functions are constructed by combining multiple kernel functions. In the present context, the generation strategy is to later use the same learning algorithm. Therefore, different settings such as learning parameters and training samples are needed. Many methods have been developed having this idea, and they can be divided into four categories. First is to manipulate the training samples. Bagging, boosting, cross-validated committees, wagging, and arcing sort of approaches are needed for this. The second way is to manipulate the input features. Methods such as random subspace, input decimation, and similarity-based feature space can be used for this purpose. The third way is to manipulate the class labels. Output coding and class switching are examples for this. The fourth way is inserting randomness to the learning algorithm. Backpropagation algorithm, randomized first-order inductive learner, and RF are few approaches used for this. The resulting classifiers using backpropagation will be quite diverse if different initial weights are applied to the same training samples in a neural network. Basically, it is visible that the underlying core of the generation strategy is to make the individual classifiers different so that it could be used to improve the classifier performance [36].

5. Machine Learning in Biomedicine

Gene expression datasets contain measurements of increasing and decreasing expression levels of a set of genes. Gene expression measurements are usually taken across time points from tissue samples or patients, and they are represented as a matrix of numerical values. In a protein-protein interaction network, the nodes represent biomolecules and the edges represent interactions. In any clustering algorithm, it is always better to have a minimum number of user inputs as it is hard for the users to specify correct values. These user inputs could affect the accuracy of the algorithm if they are incorrect [15].

Nowadays, machine learning has become ubiquitous and indispensable when it comes to solving complex problems in many areas. In biomedicine, machine learning is used for various tasks such as to predict protein structure, function from genetic sequences, discern optimal diets from patients clinical and microbiome profile, and so on. Machine learning is also used in processing real-time, high-resolution physiological data in various medical applications. The

involvement of machine learning in biomedicine can be discussed mainly in three approaches. First is that machine learning improves prognosis. Current prognostic models are restricted to few variables that humans must enter and tally scores. But these data could be directly taken from EHRs and then it would allow models to use thousands of rich predictor variables which would increase the accuracy of the predictions. Second is that machine learning will reduce the work of radiologists and anatomical pathologists. Normally, these physicians focus on interpreting digitalized images. But using machine learning algorithms, these images can be given as an input, which would result in interpretations and predictions provided by the algorithm. Sometimes, these interpretations even exceed the accuracy of humans. Apart from it, the algorithms do not need rest, and they could work continuously at the same accuracy which is not possible for humans in practice. Third is that machine learning will improve the diagnostic accuracy by minimizing the diagnostic errors. But the challenge which needs to be faced here is the complexity of training the algorithm as the predictions are not binary. Even when working with EHR data, first the data need to be preprocessed to be accessible to the algorithms as they are often stored in an unstructured format.

Data themselves are not useful, but they need to be analysed, interpreted, and acted on to make them useful. Algorithms are required to perform the aforementioned tasks with datasets. Therefore, new statistical tools from the machine learning field are critical when practicing medicine in the present time. Most of the computer-based algorithms in medicine are expert systems. They have a set of rules and knowledge which are being applied to make conclusions with regard to specific clinical scenarios. This is an approach that is similar to a medical student where the general principles about medicine are being applied on new patients. But machine learning is much different than the above discussed, as in a machine learning approach, rules are being learnt from the data themselves. It starts with the patient level observations and checks through vast number of variables to look for combinations that could reliably predict the outcome. The highlight in machine learning is the enormous number of predictors that it handles. Sometimes, there are more predictors than observations and they need to be combined in a nonlinear and highly interactive way to generate accurate outcomes.

Truly independent validation datasets must be used from different populations and periods which have no relation with the model development when testing the models. If not in the validation stage, the algorithm will have poor performance. Machine learning algorithms need high-quality data in high quantities to reach acceptable performance levels. If the dataset is biased, it can affect the performance and the generalizability of the algorithm. Machine learning is not capable of solving any fundamental problems of causal inference in the observational dataset. Even though the algorithm is good in predicting the outcomes, yet it is to be noted that these predictors are not the causes [37].

Many research studies have given binary interpretations on whether a person is being diagnosed with a particular disease or not. Sometimes, regarding a particular disease, it

may show the stage of the disease as well. It could be noticed that these research studies have been based on a particular disease, but not many diseases together. To address the said limitation, a method was proposed called ensemble label power-set pruned dataset joint decomposition (ELPPJD). The label power-set (LP) method overcomes the independence problem and takes the correlation among labels into consideration. But the issues are that the time complexity increases with the size of the label set and the imbalanced problem keeps on rising when new label sets are produced. Pruned datasets and joint decomposition can be used to overcome these issues. In this approach, when creating new classes, all these training data will be broken down into subsets which are disjoint. Then, using a similarity threshold, the labels which are more similar to each other will be grouped. The subset partition strategies can be size balanced (SB) and label similarity (LS). Random K-label sets (RAKEL) and hierarchy of multilabel classifiers (HOMER) are multilabel classifying methods which can be used as alternatives for this approach. RAKEL runs based on MEKA, and C4.5 is the basic classification algorithm. HOMER runs based on MULAN, and RF is the basic classification algorithm. RAKEL shows better performance than HOMER. But ELPPJD with LS partition strategy outperforms both RAKEL and HOMER [38].

6. Machine Learning in Biomedical Event Extraction

The relationship between the disease and the drug, the relationship between the disease and the gene, the interaction between drugs, and the interaction between proteins are biological events which have complex structures. To extract these biomedical events accurately and efficiently, biomedical text mining technology is important as the amount of unstructured and semistructured biomedical literature data is rapidly growing.

Pattern-based methods are used in biomedical relation extraction. But they are not much used in biomedical event extraction. Event extraction systems are mainly divided into two types which are rule-based event extraction systems and machine learning-based event extraction systems. In the machine learning approach, the task of extracting biomedical events is considered as a classification problem. The highly unbalanced training dataset given in biomedical event extraction is an issue, and most of the systems do not address this problem. But SVM addresses this issue using the simple class weighting strategy. Machine learning-based event extraction systems have three types. First type is the pipeline model. The pipeline model has achieved excellent results in the event extraction task. But the drawbacks in it are that the time complexity is high and each step is based on the previous step. Therefore, if there is an error, that error would be carried till the final step. Second type is the joint model. This model overcame the previous drawbacks discussed, yet it involves complicated calculations. Third type is the pairwise model. This is a combination of both pipeline and joint models. Pairwise model is faster than the joint model and more accurate than the pipeline model. This model uses

SVM to overcome the multiclass and multilabel classification problem without dealing with data imbalance.

A system proposed to extract biomedical events from imbalanced data has several steps. First, text preprocessing will be carried out using token features, sentence features, sentence dependency features, and external resource features. Next is the sample selection, based on sequential patterns. It aims to find the frequent subsequence or sequential events that satisfy the minimum support. This includes the extraction of sequential patterns in the text. Then, the detection of multiargument events is performed, followed by the joint scoring mechanism. This will result in obtaining the outcome. The tool sentence2vec, based on convolutional deep structured semantic models (C-DSSMs), is used to calculate the semantic relevance score [39].

Using SVM, it is possible to separate event extraction into multiple classification tasks, individually detecting the trigger words defining events and the arguments that describe which proteins or genes take part in these events. This is based on labelled data and supervised learning algorithms. But the issue is that these systems get affected by data sparseness. Especially, it happens when the training dataset is too small to find enough information to assign proper weights to those low-frequency or out-of-vocabulary features. Therefore, research was carried out to implement systems that would use the semisupervised or the unsupervised machine learning approach to extract biomedical events. In PubMed, a large pool of unlabelled data with potential information with regard to the biomedical event extraction domain can be found to go ahead with the research. The basic features are obtained by the labelled data and the information which is lacking in labelled data can be obtained from unlabelled data using the event feature coupling generalization strategy. Sparse features that were filtered by supervised machine learning methods can be used to increase the performance of the system while these could even be considered when using a semisupervised machine learning approach for biomedical event extraction [40].

Proteins are the end products of a gene expression. Understanding the gene function at the proteome level is a significant interest in the biological and medical research community. There are large repositories of protein data whose characteristics are unknown as noting their functional features using experimental methods have lagged far behind. This creates a need of a computational method to address the issue by accurately working on the large datasets even with a limited amount of labelled data. Functional, structural, and evolutionary characteristics of the protein can be extracted by using the protein sequence information. The aim of protein classification is to extract this information accurately, and it has only been possible with the application of machine learning algorithms to analyse the protein data repositories [41].

7. Machine Learning Approaches to Polypharmacology

Polypharmacology focuses on designing medical treatment that can target multiple receptors. The efficacy and toxicity of

drugs result from complex interactions between pharmacodynamic, pharmacokinetic, genetic, epigenetic, and environmental factors even though it may be designed as a single or multitarget treatment. Machine learning-related computational techniques are required to enable the prediction and analysis of in vitro and in vivo drug-response phenotypes.

Identification of drug-target interaction on a proteome-wide scale is essential to define and predict a drug response. Apart from its genetic and epigenetic variations, cell-to-cell communication, cell-to-cell variability, and other environmental factors must also be considered. System-based approaches help in drug discovery as they could identify cellular connectivity between all the aforementioned components. New computational methods are needed to accurately calculate free-energy landscapes during association and dissociation of protein-ligand complexes. It will allow to investigate quantitatively both low and high affinity binding on a proteome-wide scale. Self-organizing map is an unsupervised machine learning technique that is used to cluster drug molecules. Prediction of drug-target interactions can be done by combining the sequence features of the receptor with the fingerprint of the ligand, in order to train modules based on statistical machine learning. There are four computational areas which are crucial in polypharmacology, and they are the proteome-wide prediction of drug-target interactions, the quantitative modelling of protein-ligand interactions, the integrated analysis of biological networks, and the dynamic and stoichiometric simulation of biological networks [42].

NGS includes both DNA and RNA sequencing. It involves the task of breaking DNA and RNA into fragments and determining the order of the nucleotide bases in each fragment. DNA sequencing can either be whole genome sequencing (WGS) or whole exome sequencing (WES) of the coding regions of all known genes and targeted sequencing of genomic regions or genes implicated in a disease. WGS and WES can be found in clinical practices to evaluate developmental brain disorders such as autism, seizures, and intellectual disability. There are commercially available sequencing platforms which use various methods to generate sequence data. With the continuous modifications resulted from the advancement of technologies, NGS-based platforms have shown a comparatively low error rate. In NGS technologies, Sanger sequencing is considered to be the most accurate sequencing method and it is often used for the validation of genetic variants due to its high accuracy [43].

When modern machine learning is combined with medical practice by using clinical data sources, various prediction models could be generated for similar clinical questions. From early warning systems till superhuman imaging diagnostics, there are many applications of machine learning in medical practice. Machine learning methods provide predictions based on existing data. The Google flu reminds how forecasting an annual event based on 1 year of data could lead to time series problems. Having loads of data over the past time has diminishing returns. Basically, it is found in research on decision support algorithms that it is better to use the most recent year data instead of historical

data of multiple years for better accurate decisions. Even when evaluating prediction models, the main idea behind this is not about the ability to do repetitions of historical trends but the accuracy it has on predicting future trends. It is true that machine learning algorithms can improve the prediction accuracy by using conventional regression models by capturing complex and nonlinear relationships in the data. But it is to be noted that even though highest computing power is provided, there is no possibility of getting information that is not present. Therefore, using clinical data alone has limited the prediction power. By incorporating the correct data streams, the level of prediction accuracy can be improved. But not even a simple nonlinear system can precisely predict the distant future. Tiny variations which seem nothing and very less important rounding errors can accumulate and make a massive impact on future events [44].

8. Machine Learning for Drug Repurposing Using System Biology

More than 90% of the drugs that go through the early phases of clinical trials fail due to reasons such as adverse reactions, side effects, or lack of efficiency. To overcome these challenges, drug repurposing has been considered. Repurposing drugs can be either drug-based or disease-based. Drugs that are strongly anticorrelated with a disease are likely to be candidates for repurposing. The connectivity map was the first project that aimed to explore the functional connectivity between drugs. It even considered the functional connectivity between drugs and diseases. Systems biology can be used to discover and develop drugs by giving concern to the interactions of the components in biological entities. Drugs are being ranked based on the amount of perturbation they cause on specific disease-related genes.

A drug disease network (DDN) is constructed by integrating the knowledge on disease-related genes, drug targets, signalling pathways, and gene-gene interactions. The DDN represents all the interactions between drug targets and genes, related to a given disease as described in Kyoto Encyclopaedia of Genes and Genomes (KEGG) signalling pathways. The repurposing scores of the drug-disease pairs are calculated using the Pearson correlation coefficient between their gene perturbation signatures, and it can take a value between 1 to -1 . A high positive value indicates that the drug and the disease both cause similar perturbations to the system, and a high negative value indicates that the drug and the disease have opposite gene perturbation signatures. By using this value, it can be recognized whether it is a drug that could be considered as a treatment to a given disease [45].

Therefore, it is clear that when a patient approaches and expects support from healthcare, the decision making starts in the process of the patient being diagnosed with a disease and it continues till the proper treatment or medication is prescribed for the patient. In this process, which is a chain of decisions, to support each decision, machine learning approaches are involved in healthcare. Tasks such as predicting or diagnosing a disease, identifying hidden diseases,

providing clinical decision support, and even recognizing whether a drug is suitable as a treatment to the given disease are being conducted using machine learning approaches to support clinicians to make fast and accurate decisions. Stepping further ahead, even after a patient is cured from a disease, using machine learning approaches, the EHRs of patients are processed and analysed to identify any future health risks that could possibly occur.

9. Discussion

In this section, the main concern would be given towards highlighting important and crucial facts of the topics discussed throughout the paper. To start with, it is better to give concern towards the performance of machine learning algorithms. To identify best performing algorithms, classifier accuracy and the classifier log loss are two factors that can be considered. The classifier accuracy needs to be high and the classifier log loss needs to be low for an algorithm to be identified as a well-performing algorithm. Therefore, once selecting a suitable algorithm to address a specific concern, the aforementioned factors are considered to select the algorithm out of many different existing algorithms that would best suit our purpose.

In most clustering algorithms, the effectiveness of the algorithm depends on the appropriate parameters that the user inputs. For example, parameters like number of expected clusters, a starting point in the dataset, minimum number of samples to form a cluster, and so on affect the clustering result. This is a serious issue when biological and biomedical data are considered as they are nonspherical and high-dimensional. Automatic density clustering method with multiple kernels (ADCMK) is a proposed option to answer the said issue. Using that method, clusters with arbitrary shape can be easily identified based on their density. When comparing the conventional clustering algorithms, ADCMK automatically determines optimal values for the cutoff distance, kernel weights, and number of clusters and centroids which will not change the clustering result when executed time to time. It also has comparatively better accuracy. Multiple kernel clustering approaches help in optimally combining the learning algorithms in order to obtain excellent clustering results and performance in various scenarios. They are mostly suitable to handle labelled datasets as they are a part of supervised kernel learning. There is also an unsupervised multiple kernel learning approach that aims to determine the linear combination of multiple kernels based on an unlabelled dataset [46].

It is not possible to exactly highlight which algorithm is better than the other. The reason is that it depends on the domain that the training and the tests are being executed, the dataset that is involved in the training and testing procedures, the level of preprocessing that is being performed on the dataset, the selected feature set for the algorithm or the feature selection algorithms that have been used on the dataset, the size of the dataset and the data types in the dataset, the performance level and the capacity of the machine, and much more. This makes it vital to select the proper algorithm that would be ideal for the requirement.

Normally, this cannot be directly selected, but it will be an iterative process to select the most suitable algorithm from a filtered set of algorithms which would best fit the problem in hand. The proper knowledge and past experiences gathered will help in constructing the filtered set of algorithms. When looking at a simple example from the biomedical field, it can be seen that for the diagnosis of various diseases, the best performing algorithm changes due to the aforementioned reason. This could be further explained by looking at the scenarios below, in detail.

Many studies have been carried out using ANN, DT, and LR to diagnose kidney diseases, and ANN has outperformed both DT and LR by a huge margin [26]. According to previous studies, it is to be noted that the accuracy of the results obtained with regard to lung cancer diagnosis is in the order of SVM, ANN, and DT. Apart from them, various other techniques also have been used and they have also showed a reasonable level of accuracy. These algorithms are GA, BPNN, Hopfield neural network, and LDA [30]. Another benefit that could be gained by using SVM classification is that through SVM, it is possible to calculate even the stage of the lung cancer [18]. Regression can be used to model and predict the association between the dependent variable and the independent variable. Regression can be categorized as linear regression and logistic regression. For linear regression, the dependent variable should be continuous, and for logistic regression, the dependent variable should be discrete. Using the logistic regression algorithm to predict complex disease such as cancer must be avoided. KNN is not suitable for cancer predictions as the datasets are huge and more complex which could make the clustering process more difficult [47]. It was found that DNN has better performance for breast cancer detection. Then ANN, SVM, and KNN were ranked in the given order with regard to their performance [48]. It is to be noted that DNN has also shown higher accuracy in predicting diabetics compared with the other standard machine learning algorithms. When discussing extracting of medical data from documents, datasets, databases, medical images, etc. for analysing purposes, various CNN algorithms can be used to extract features from unstructured data. For structured data, NB, KNN, and DT algorithms can be used. For textual data which are unstructured, CNN-based unimodal disease risk prediction (CNN-UDRP) can be used. For both structured and unstructured data, CNN-based multimodal disease risk prediction (CNN-MDRP) can be used. In the CNN-UDRP algorithm, tests are represented in the form of a vector. For each word, the distributed representation of the word embedding found in NLP is used. When working with CNN, it is required to specify the number of words for the sliding window. When assigning a value for the sliding window, it is important to assign a value that would help in reaching the highest accuracy when it is being used [8]. When giving concern to these aspects, what could be seen is how these machine learning approaches assist in healthcare decision making of the clinicians. When analysing reports and when going

through the patients' medical history, these algorithms are used to highlight the factors that definitely need to be considered by the clinician when prescribing treatment. In addition to it, these algorithms are also used to assist in identifying the level that the disease has spread in the patient so that the treatment can be given accordingly based on the disease severity.

Moving on to the next topic, which is medical imaging, it is visible that currently, medical image classification is mostly based on pattern recognition methods. It has been found that classification-based neural networks have better performance than other supervised machine learning algorithms. Deep learning is considered as a combination of both supervised and unsupervised methods. The reason is that it initially relies on unsupervised learning for training the DNN and then to fine tune the neural network, it uses supervised learning. When compared with various classification algorithms, CNN combined with adaptive sliding window fusion mechanism has good stability and robustness along with high classification accuracy. Even when it is for breast tumour or brain tumour diagnosis, the aforementioned approach with CNN has given comparatively better results [49]. As both deep learning and ensemble learning can construct complicated nonlinear functions, by combining them, it is possible to handle more complex tasks. Biomedical time series often has interfering noise. But when trying to eliminate the noise, along with it some useful information might also be lost. DNN has the ability to capture useful information while ignoring the interfering noise after learning from training samples. Presently, DNN mainly includes CNN, deep belief network (DBN), and autoencoder (AE). CNN performs convolution operations on both horizontal and vertical directions. This is a good approach for image data as it is relevant in both directions. But for biomedical time series with multiple channels, this is not suitable as only horizontal direction is relevant and the vertical direction is independent. For this reason, multi-channel CNN can also be implemented to obtain better classification performance [36].

When further observed, it is also visible that in the present context, medical imaging has also been combined with the 3D printing technology to achieve greater success specially when surgical planning is to be considered. 3D printing can turn a 3D medical image dataset into a 3D object as per the patient's actual size. This can largely improve the precision of modern surgical planning by combining with other preoperative plan and intraoperative navigation techniques. By combining these with the cloud computing architecture, there is further space for improvement in cost efficiency, flexibility, and clinical workflow. Therefore, further exploration needs to be done on 3D printing and cloud computing technologies of their usage in the surgical realm. Nowadays, surgeons are able to perform risky procedures with high level of safety and accuracy with the development of multimodality imaging, with regard to especially the fusion of functional imaging in preoperative planning and modern image guided therapy in intraoperative navigation. With the development of medical imaging technology, there are multiple imaging processing

systems and intraoperative imaging guided platforms which are being introduced that are powerful but also complicated [50]. In addition to it, 3D printing has also been used to design advanced prosthetic devices with the aim of developing a realistic model which is closely similar to the actual part with regard to the geometry and mechanical properties. When making such products, even though the geometry can be easily obtained by the 3D scanned data, the mechanical properties need to be mechanically validated and a major concern should be given towards composite material technology [51]. A few of the currently existing findings on prosthetic devices are briefly discussed below.

The concept of creating a soft total artificial heart from silicon elastomers using the lost-wax casting technique has also been possible as the 3D printing technology has enabled the production of entirely soft pumps with complex chamber geometries. Even though it is not yet readily implantable, it promotes the future of personalized medicine and artificial heart development without requiring expensive and complex equipment [52]. By using an array of semiconductor photo detectors which are made of polymers and printed on a glass hemisphere, a 3D printed prototype of a bionic eye has been designed and further research is being carried out on ways to improve the efficiency of the photo detectors and to find a way to print on a soft hemispherical material that could be implanted to a real human eye [53]. Researchers have also developed a 3D printed tooth that also kills bacteria using conventional artificial tooth resins combined with positively charged quaternary ammonium ions [54]. In another study, a handheld portable printer was developed which could be used to create narrow sheets of skin tissue made from collagen and fibrin to cover and heal deep wounds. This printer can tailor tissues based on the patient and the wound characteristics within two minutes or less. Further research is being carried out on ways to expand the size of the printed strips so that larger wounds could be covered [55]. Bioengineers have 3D printed artificial hyperelastic bones which could quickly induce bone regeneration and growth. Currently, it has been implanted successfully in animal models and a main advantage in this approach is that the 3D printing process can customize any shape of a bone as per the requirement and can create it within a day. It is to be noted that the materials used to create these 3D printed bones are flexible, and after combining them with the bone structure, with time, they bond with the existing bones and become a part of the human body [56]. A group of scientists has 3D printed a bionic ear with an antenna and a cartilage that can hear the radio frequencies even beyond the range of a normal human's capability. It has also been created in such a way that the electrical signals produced from the ear can be connected with the nerve endings which is considered to be the first time that 3D printing was effectively used to interweave tissues with electronics [57]. There are many such prosthetic devices which were outcomes of previous research studies that were carried out, and further research is being conducted on them presently to improve the existing results. Medical imaging, 3D printing technologies, and various other computational approaches have been vital when conducting these research

studies to obtain improved results. When giving concern to how machine learning could be involved in these findings, it is visible that machine learning approaches could analyse large biological and material datasets to identify various relationships which exist between the data. In addition to it, as machine learning is used to detect diseases and the severity level of it, these machine learning algorithms can be combined with the equipment that is being used to design such prosthetic devices so that when the medical dataset is given, an analysis is to be performed so that the devices can be customized and designed based on the patient's requirement which would further encourage the latest trend in healthcare, which is personalized medicine.

When discussing biomedicine, biomedical data play an important role which cannot be neglected. It is vital to extract the required knowledge from biomedical data that would make paths visible to explore the subject matters further. When giving concern towards bioinformatics and biomedical data, there are various documents that contain high-dimensional data with large number of attributes such as categorical images and text records. To extract information from such documents efficiently and with high accuracy, t-distributed stochastic neighbour embedding method could be used. It is a popular method for exploring high-dimensional data in the field of machine learning and can be incorporated with the multiple kernel approach as well. ADCMK performs data preprocessing, optimal combined kernel determination, cluster centroid detection, and assignment and visualization of clusters which shows excellent results when working with larger-scale biological and biomedical datasets without class label information [46].

Under biomedical event extraction, the challenges that protein classification tasks need to face due to the limited data that are available for learning and the unbalanced classes in those existing data need to be considered and addressed. Expectation-maximization-based algorithms are there to address these challenges by incorporating unlabelled data. Even by using the n-gram model for the feature space with the NB classifier, the accuracy of the model could be improved to some extent without using unlabelled data. The transductive support vector machine (TSVM) method is a semisupervised learning approach that also could be used in addressing the said issues. But there are other semi-supervised algorithms as well that could perform better in certain classification problems [41].

While coming towards the machine learning approaches to polypharmacology, it is important to recognize about certain risks or loopholes which can be found when integrating machine learning algorithms with biology and biomedicine. In some occasions, there may be identical twins with the same observable demographic characteristics, lifestyle, medical care, and genetics which will provide same predictions by a machine learning algorithm but will end up in totally unexpected different outcomes. This is an instance to show that the incorporation of machine learning with biology and biomedicine should be performed with maximum care and monitoring. Research studies are growing in number trying to increase the accuracy of the predictions made,

but still these algorithms do not state how to change the obtained outcome. It even does not assure whether it is possible to change the received outcomes in any possible way. Machine learning-based identifications are strong, but they are theory free. This means that no theory-based explanations can be provided, even for the generated predictions. Sometimes, the most accurate predictions made by machine learning algorithms are obvious to the practicing clinicians as well. Even though predictive algorithms cannot eliminate medical uncertainty, they improve allocation of scarce healthcare resources. Using DNN, now early warning systems can be rapidly developed and optimized using real-world data. These systems will also have capabilities which seemed to be impossible to be achieved previously [44].

Here onwards the discussion will give concern towards increasing the productivity of applications through increasing the performance of the machine learning algorithms. In order to increase the performance and the efficiency of machine learning algorithms, various steps and measures are being carried out and tested by the researchers. When observing the said fact, next are a few instances to get an idea of what sort of tasks are being conducted to increase the performance of the algorithms. When considering the accuracy, sensitivity, and the specificity of classifiers while changing the number of folds, any major fluctuation was not found in the values obtained in the confusion matrix. All were just small variations in the final decimal places which can be ignored [26]. But DNN, feature selection methods, and cross validation techniques can be applied to increase the classification accuracy [10, 58]. It is also to be noted that positive predictive value (PPV), negative predictive value (NPV), accuracy, sensitivity, and specificity attributes show an improvement when a classifier is applied after performing PCA rather than the classifier being applied to the official feature set (OFS). But when the GA was used instead of the PCA, all the above attributes showed a further improvement even than what was resulted after the PCA has been applied [29].

Ensemble learner is constructed by a collection of individual classifiers. When there is a pattern classification problem, it is solved by a combination of multiple classifier outputs. Majority voting, bagging, and boosting are few well-known ensemble methods which can be used to combine the outputs. Normally, this approach is more accurate than using a single classifier. In the majority voting method, the label to be assigned is decided based on the majority of the individual classifier's outputs. The bagging method decreases the variance of the predictions by generating additional data for training from the original dataset. The boosting method has two steps. The first is using the original data to produce a series of averagely performing models. Then, the next step is to increase the performance by combining the previous performances using a voting scheme [10]. Using the ensemble technique, a weak classifier can be converted to a strong classifier. When working with a weak classifier, it is merged using the weighted average of the prediction accuracy as metrics. Then, the error will be propagated

with every prediction, and until the prediction becomes accurate, multiple iterations will be carried out to reduce the error [59].

Maximum sensitivity neural networks use the back-propagation algorithm to adjust the weights associated with the neurons. This is special as it saves time and memory while predicting more accurately as the output layer detects the maximum sensitivity of the neurons for the new pattern [47].

Machine learning feature extraction is the process that helps in identifying the relevant attributes from various candidate subnets, and it plays a vital role in creating an effective predictive model. There are various benefits of applying feature selection, such as highly effective and fast training of the machine learning algorithm, reducing the complexity of the model, improving the accuracy of the model, and eliminating the overfitting problem. There are three groups that could be identified in feature selection methods. They are filter, wrapper, and embedded methods. The filter method is normally a preprocessing step that relies on general features. The wrapper method uses machine learning in order to select the best and the most suited subset of features. Forward feature selection, backward feature elimination, and recursive feature elimination are widely used wrapper methods. The embedded method combines the qualities of the filter and wrapper methods [17]. It is important to remember that when features are correlated to other features, there is no point having all of them considered in the machine learning algorithm as it will increase only the execution time but will not improve the efficiency of the algorithm. It is also to be noted that when adding certain features, the accuracy of the algorithm may drop. It is preferred if each of the features could be tested individually and then combined in a way that the accuracy of the output will be increased [27]. The quality of the feature selection process has a direct impact on the accuracy increase. Some algorithms like ANN show a higher increase in the accuracy when compared with other algorithms after the feature selection phase is performed [25].

Machine learning techniques have gained significantly high accuracy when considering classification-based problems. A four-tier architecture can be used to store and process a huge volume of data efficiently. Tier 1 can be used to collect data, and tier 2 can be used to store huge volumes of data. Tier 3 and tier 4 can be used for machine learning classifications and representation of the results, respectively [14]. In the present time, as computing power is not an issue, DNN has more than 20 layers. Early neural networks were typically less than 5 layers. DNNs are used for various purposes such as automatic object detection, segmentation on images, automatic speech recognition, genotypic and phenotypic detection, classification of diseases in bioinformatics, and so on. Even though CNN is quite similar to regular neural networks, it assumes that there are geometric relationships among inputs. By using convolutional layers in DNN, important features could be amplified for better analysis. The pooling layer in CNN takes the output and finds the maximal value so that it could give the convolution function the best extracts of the images. Regularization is

used when training deep networks as it increases the performance by reducing overfitting. There are no exact number of layers to obtain more accurate results, but it is an iterative trial and error process in which the best architecture needs to be found for a specific problem. The advantage of using CNN over traditional machine learning algorithms is that there is no need to compute the features as the first step. But CNN will find the important features as a part of its searching process [8].

As ANN is complex, other simple machine learning algorithms such as SVM, KNN, and RF gradually came ahead of ANN in popularity. By selecting the appropriate activation functions, better feature extraction can be achieved, as they are what form the nonlinear layers in deep learning frameworks. There are many activation functions. Sigmoid function, hyperbolic tangent, softmax, rectified linear unit (ReLU), softplus, absolute value rectification, and maxout are a few of such commonly used activation functions. There are various deep learning architectures such as AE, DBN, restricted Boltzmann machines (RBM), CNN, and RNN. AE is different from ordinary ANN as it extracts features from unlabelled data and sets target values to be equal to the inputs. RBMs are generative graphical models and their aim is to learn the distribution of the training data. DBM can be considered as a stack of AEs or RBMs. CNN is different from other deep learning algorithms as it extracts features from small sections of an input image. RNN outperforms all the other deep learning approaches when it is about dealing with sequential data. Machine learning is widely used in the medical image domain. Sparse autoencoder (SAE) and RBM can deal with dynamic data and they can also extract patterns from unlabelled data. Due to its capacity of analysing special data, CNN is used in biomedical image analysis commonly. When considering the challenges, it is to note that in deep learning, it requires a large amount of labelled data for model training. Another issue when working with biomedical data is that they are imbalanced, as the normal people set is much larger than the diagnosed people set. They also need high memory and high computing power. The input images must be in high resolution. Nowadays, there are various medical devices and sensors which provide large amounts of data. Deep learning approaches are good interpreters of these data for disease prediction, prevention, and diagnosis. Table 1 lists few applications of the above-discussed deep learning algorithms in the field of biology and biomedicine [60].

As per Table 2, there are several open-source machine learning libraries in various languages which could be used to assist in the development process when implementing intelligent biological applications. Python and C++ have the most number of libraries that support deep learning. But Python is more often used in traditional machine learning approaches [60]. The popularity or the usage of various deep learning frameworks in projects, when categorized based on the language that they are written in, is shown in Figure 1.

At the initial stage, simple machine learning classifying and clustering algorithms were used in data processing and predictions, but when moving forward, with the development in computational power and programming

TABLE 1: Applications of deep learning algorithms in computational biology [60].

Deep learning algorithm	Medical image analysis	Protein structure prediction	Genomic sequencing and gene expression analysis
Convolutional neural network	Brain tumour segmentation, knee cartilage segmentation, prediction of semantic descriptions from medical images, segmentation of MR brain images, coronary artery calcium scoring in CT images	Prediction of protein order/disorder regions, prediction of protein secondary structures, prediction of protein structure properties	—
Sparse autoencoder	Organ detection in 4D patient data, segmentation of hippocampus from infant brains, histological characterization of healthy skin and healing wounds	Sequence-based prediction of backbone $C\alpha$ angles and dihedrals	—
Deep belief network	Segmentation of left ventricle of the heart from MR data, discrimination of retinal-based diseases	Prediction of protein disorder, prediction of secondary structures, local backbone angles	Modelling structural binding preferences and predicting binding sites of RNA-binding proteins, prediction of splice junction at DNA level
Deep neural network	Brain tumour segmentation in MR images, prostate MR segmentation, gland instance segmentation	—	Gene expression inference, prediction of enhancer, prediction of splicing patterns in individual tissues and differences in splicing patterns across tissues
Recurrent neural network	Classification of patterns of EEG synchronization for seizure prediction, EEG-based lapse detection	Prediction of protein secondary structure, prediction of protein contact map	Prediction of miRNA precursor and miRNA targets, detection of splice junctions from DNA sequences

TABLE 2: Machine learning libraries in languages [8].

Language	Traditional machine learning libraries	Deep neural network machine learning libraries
Python	Scikit-learn, PyBrain, Nilearn, Pattern, MILK, Mixtend	Keras, Tensorflow (written in both C++ and Python), Nolearn, DeePy, Pylearn2
R	Caret, Boruta, GMMBoost, H2O, KLaR, rminer	Darch, DeepNet
C++	Shogun	Caffe, EBLearn, Intel Deep learning Framework, Tensorflow (written in both C++ and Python)
Java	Encog, Spark, Mahout, MALLET, Weka	Deeplearning4j
JavaScript	Cluster, LDA, Node-SVM	ConvnetJS

methodologies, more enhanced and complex machine learning algorithms were adapted by the field of biology and biomedicine to obtain more accurate results at a higher efficiency. Even at this very moment various attempts are being executed to further evolve the computational biology and biomedicine domain with the integration of sophisticated scalable machine learning algorithms. Many languages provide multiple application program interfaces (APIs) and libraries in order to assist in these development processes so that the main concern or priority could be given to the requirement with regard to the subject matter and less attention towards the efficiency and performance of the algorithms, as they have already been handled by the libraries or APIs. These algorithms could even be applied to different domains such as marketing, social media, search engines, fraud detection, stock market analysis, e-business, authentication, and so on by changing the training datasets and the features to be analysed. But in this review paper, the computational biology and biomedicine domain was considered.

When coming back to precision medicine, as discussed previously, the simplest definition that could be provided would be treating patients based on their genetic, lifestyle, and environmental data. This is simply not an easy task. It needs to have a whole load of data that should be analysed to obtain the required knowledge for it. This is near impossible for humans to manually process. Furthermore, some relationships among data cannot be identified by humans after exploring the dataset. But as discussed previously, unsupervised machine learning has been able to capture these relationships and group data according to the identified relationships. As precision medicine studies are involved with genes, it is certain that the biomedical event extraction would also contribute towards it. For a patient to then be treated using precision medicine, few approaches in computational biology are vital, such as biomedical data analysis, disease diagnosis, and medical image analysis. The reason is that categorizing the patients to identify the treatment to be followed according to a set of considered factors needs better observations of the patient and his/her medical history.

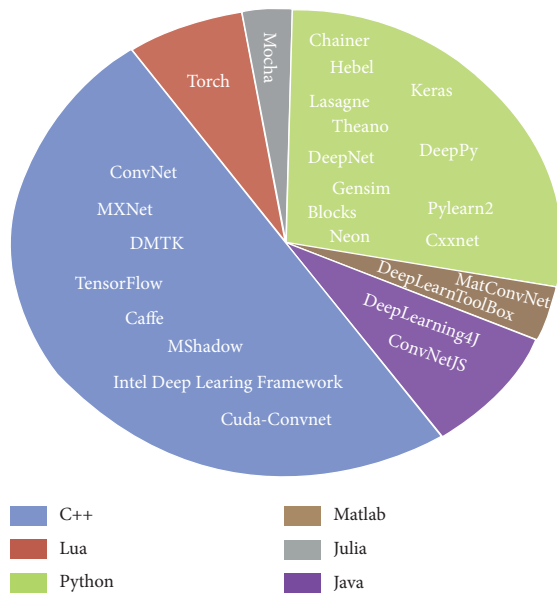


FIGURE 1: Popularity of deep learning frameworks [60].

Throughout this paper, on many occasions, the involvement of machine learning in computational biology and biomedicine has been discussed. The reason behind it is that more than 80% of the healthcare decisions are made based on the said areas. In addition to it, data analytics is being applied in the healthcare sector to improve the healthcare services by predicting the future expectations based on a patient's medical history. Therefore, there is no doubt that prominent attention needs to be given towards the areas of computational biology and biomedicine when discussing computational decision making in healthcare. But it is to be highlighted that whenever we use an approach related to artificial intelligence in healthcare, it is always related to making a decision. Disease detection, disease prediction, drug repurposing, precision medicine, medical resource allocation, and much more can be shown as evidence for it. In all of them, a machine learning algorithm either makes a decision or least it would support a decision. It helps in solving a doubt that a clinician may have or sometimes highlighting a fact which was not visible to the clinician at the time of making a decision as a solution for a healthcare-related issue. In addition to it, most importantly, it deals with people's lives and even a simple mistake can make a huge impact on a person and even the society at large. Therefore, computational decision making is so crucial in the healthcare sector to make confident and accurate decisions in less time and at low cost. This has only been possible with machine learning approaches being adapted as a tool in healthcare and is currently being extensively applied in healthcare decisions.

10. Conclusion

While reaching the end of the paper, there is no doubt that machine learning algorithms, which is a subsection of artificial intelligence, have inspired the field of computational biology and have contributed immensely to the healthcare

sector in terms of fast, efficient, accurate, and cost-effective computational decision making. It is clearly visible that in the present context, machine learning has been applied in various sections in the discussed field. The involvement of machine learning can be found in disease diagnosis and prediction, medical imaging, drug repurposing, biomedical event extraction, and much more in healthcare. But what is certain is that the journey that started with the integration of machine learning to computational biology has come a long way passing several milestones and is now at a peak with the introduction of precision medicine. When considering all the above applications of machine learning in healthcare, what is clearly visible is that how artificial intelligence has been a key resource area for decision making in the healthcare sector in various aspects. When giving concern to the most recent applications of artificial intelligence in healthcare, the best example would be how different tasks such as patient care, treatment research, resource allocations based on the hospital volume, predictions and preparing for future possible requirement, and so on were handled and managed during the COVID-19 pandemic. With all that being said, it is crystal clear that artificial intelligence had played a huge role through machine learning to implement computational decision making tools for the healthcare sector, and in present times, they are of much importance and could not be separated from the healthcare sector.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to specially thank the respective staff members in the Faculty of Information Technology of University of Moratuwa who guided and assisted them to write this paper.

References

- [1] S. Sánchez Martínez, O. Camara, G. Piella et al., "Machine learning for clinical decision-making: challenges and opportunities," 2019.
- [2] S. Debnath, D. P. Barnaby, K. Coppa et al., "Machine learning to assist clinical decision-making during the COVID-19 pandemic," *BioMed Central*, vol. 6, pp. 2332–8886, 2020.
- [3] T. Lysaght, H. Lim, V. Xafis, and K. Ngiam, "AI-assisted decision-making in healthcare," *The Application of an Ethics Framework for Big Data in Health and Research*, vol. 11, pp. 299–314, 2019.
- [4] A. Calamuneri, L. Donato, C. Scimone, A. Costa, R. D'Angelo, and A. Sidoti, "On machine learning in biomedicine," *Life Safety and Security*, vol. 5, no. 12, pp. 96–99, 2017.
- [5] A. Bharat, N. Pooja, and R. A. Reddy, "Using machine learning algorithms for breast cancer risk prediction and diagnosis," in *Proceedings of the 3rd International Conference on Circuits, Control, Communication and Computing*, Bangalore, India, July 2018.
- [6] R. C. Deo, "Machine learning in medicine," *Circulation*, vol. 132, no. 20, pp. 1920–1930, 2015.

- [7] V. Mishra, Y. Singh, and S. Kumar Rath, "Breast cancer detection from thermograms using feature extraction and machine learning techniques," in *Proceedings of the IEEE 5th International Conference for Convergence in Technology*, Bombay, India, March 2019.
- [8] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine learning for medical imaging," *RadioGraphics*, vol. 37, no. 2, pp. 505–515, 2017.
- [9] P. Radhika, R. Nair, and G. Veena, "A comparative study of lung cancer detection using machine learning algorithms," in *Proceedings of the IEEE International Conference on Electrical, Computer and Communication Technologies*, Coimbatore, India, November 2019.
- [10] A. Al-Zebari and A. Sengur, "Performance comparison of machine learning techniques on diabetes disease detection," in *Proceedings of the 1st International Informatics and Software Engineering Conference*, Ankara, Turkey, November 2019.
- [11] M. R. Al-Hadidi, A. Alarabeyyat, and M. Alhanahnah, "Breast cancer detection using K-nearest neighbor machine learning algorithm," in *Proceedings of the 9th International Conference on Developments in eSystems Engineering*, Liverpool, UK, September 2016.
- [12] M. S. Yarabarla, L. K. Ravi, and A. Sivasangari, "Breast cancer prediction via machine learning," in *Proceedings of the 3rd International Conference on Trends in Electronics and Informatics*, Tirunelveli, India, April 2019.
- [13] S. Sharma, A. Aggarwal, and T. Choudhury, "Breast cancer detection using machine learning algorithms," in *Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems*, Belgaum, India, June 2018.
- [14] M. R. Ahmed, S. M. Hasan Mahmud, M. A. Hossin, H. Jahan, and S. R. Haider Noori, "A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms," in *Proceedings of the IEEE 4th International Conference on Computer and Communications*, Chengdu, China, April 2018.
- [15] B. Andreopoulos, A. An, X. Wang, and M. Schroeder, "A roadmap of clustering algorithms: finding a match for a biomedical application," *Briefings in Bioinformatics*, vol. 10, no. 3, pp. 297–314, 2009.
- [16] J. Latif, C. Xiao, A. Imran, and S. Tu, "Medical imaging using machine learning and deep learning algorithms: a review," in *Proceedings of the 2nd International Conference on Computing, Mathematics and Engineering Technologies*, Sukkur, Pakistan, March 2019.
- [17] H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, and M. Faisal Nagi, "Automated breast cancer diagnosis based on machine learning algorithms," *Journal of Healthcare Engineering*, vol. 2019, Article ID 4253641, 2019.
- [18] W. Rahane, H. Dalvi, Y. Magar, A. Kalane, and S. Jondhale, "Lung cancer detection using image processing and machine learning healthcare," in *Proceedings of the International Conference on Current Trends towards Converging Technologies*, Coimbatore, India, March 2018.
- [19] Q. Wu and W. Zhao, "Small-cell lung cancer detection using a supervised machine learning algorithm," in *Proceedings of the International Symposium on Computer Science and Intelligent Controls*, Budapest, Hungary, October 2017.
- [20] S. Mandal and V. Daivajna, "Machine learning based system for automatic detection of leukemia cancer cell," in *Proceedings of the IEEE 16th India Council International Conference*, Rajkot, India, December 2019.
- [21] S. Mani, Y. Chen, T. Elasy, W. Clayton, and J. Denny, "Type 2 diabetes risk forecasting from EMR data using machine learning," pp. 606–615.
- [22] S. Priya and R. Rajalaxmi, "An improved data mining model to predict the occurrence of type-2 diabetes using neural network," in *Proceedings of the International Conference on Recent Trends in Computational Methods, Communication and Controls*, Kochi, India, August 2012.
- [23] S. Ayon and M. Islam, "Diabetes prediction: a deep learning approach," *International Journal of Information Engineering and Electronic Business*, vol. 7, no. 6, pp. 21–27, 2019.
- [24] S. Niazi, H. A. Khattak, Z. Ameer, M. Afzal, and W. A. Khan, "Cardiovascular care in the era of machine learning enabled personalized medicine," in *Proceedings of the International Conference on Information Networking*, Barcelona, Spain, April 2020.
- [25] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mobile Information Systems*, vol. 2018, Article ID 3860146, 2018.
- [26] Y. Amirgaliyev, S. Shamilulu, and A. Serek, "Analysis of chronic kidney disease dataset by applying machine learning methods," in *Proceedings of the IEEE 12th International Conference on Application of Information and Communication Technologies*, Almaty, Kazakhstan, October 2018.
- [27] S. Shetty and Y. S. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis," in *Proceedings of the International Conference on Inventive Computation Technologies*, Coimbatore, India, August 2016.
- [28] S. Aich, K. Younga, K. L. Hui, A. A. Al-Absi, and M. Sain, "A nonlinear decision tree based classification approach to predict the Parkinson's disease using different feature sets of voice data," in *Proceedings of the 20th International Conference on Advanced Communication Technology*, Chuncheon-si Gangwon-do, Korea (South), December 2018.
- [29] S. Aich, H. Kim, K. Younga, K. L. Hui, A. A. Al-Absi, and M. Sain, "A supervised machine learning approach using different feature selection techniques on voice datasets for prediction of Parkinson's disease," in *Proceedings of the 21st International Conference on Advanced Communication Technolog*, PyeongChang Kwangwoon_Do, Korea (South), February 2019.
- [30] Ö. Günaydin, M. Günay, and Ö. Sengel, "Comparison of lung cancer detection algorithms," in *Proceedings of the Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science*, Istanbul, Turkey, December 2019.
- [31] Y. Hasija, N. Garg, and S. Sourav, "Automated detection of dermatological disorders through image-processing and machine learning," in *Proceedings of the International Conference on Intelligent Sustainable Systems*, Palladam, India, December 2017.
- [32] D. Solari et al., "Skull base reconstruction after endoscopic endonasal surgery: new strategies for raising the dam," in *Proceedings of the II Workshop on Metrology for Industry 4.0 and IoT*, Naples, Italy, June 2019.
- [33] F. Bauer, *Imaging and Diagnosis for Planning the Surgical Procedure*, IntechOpen, London, UK, 2020.
- [34] R. C. Miner, "Image-guided neurosurgery," *Journal of Medical Imaging and Radiation Sciences*, vol. 48, no. 4, pp. 328–335, 2017.
- [35] R. Pratt, J. Deprest, T. Vercauteren, S. Ourselin, and A. L. David, "Computer-assisted surgical planning and intraoperative guidance in fetal surgery: a systematic review," *Prenatal Diagnosis*, vol. 35, no. 12, pp. 1159–1166, 2015.

- [36] L.-p. Jin and J. Dong, "Ensemble deep learning for biomedical time series classification," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 6212684, 2016.
- [37] Z. Obermeyer and E. J. Emanuel, "Predicting the future—big data, machine learning, and clinical medicine," *New England Journal of Medicine*, vol. 375, no. 13, pp. 1216–1219, 2016.
- [38] R. Li, W. Liu, Y. Lin, H. Zhao, and C. Zhang, "An ensemble multilabel classification for disease risk prediction," *Journal of Healthcare Engineering*, vol. 2017, Article ID 8051673, 2017.
- [39] Y. Lu, X. Ma, Y. Lu, Y. Zhou, and Z. Pei, "A novel sample selection strategy for imbalanced data of biomedical event extraction with joint scoring mechanism," *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 7536494, 2016.
- [40] J. Wang, Q. Xu, H. Lin, Z. Yang, and Y. Li, "Semi-supervised method for biomedical event extraction," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, Philadelphia, PA, USA, October 2012.
- [41] B. R. King and C. Guda, "Semi-supervised learning for classification of protein sequence data," *Scientific Programming*, vol. 16, no. 1, pp. 5–29, 2008.
- [42] L. Xie, L. Xie, S. L. Kinnings, and P. E. Bourne, "Novel computational approaches to polypharmacology as a means to define responses to individual drugs," *Annual Review of Pharmacology and Toxicology*, vol. 52, pp. 361–379, 2011.
- [43] T. C. Carter and M. M. He, "Challenges of identifying clinically actionable genetic variants for precision medicine," *Journal of Healthcare Engineering*, vol. 2016, Article ID 3617572, 2016.
- [44] J. H. Chen and S. M. Asch, "Machine learning and prediction in medicine—beyond the peak of inflated expectations," *New England Journal of Medicine*, vol. 376, no. 26, pp. 2507–2509, 2017.
- [45] A. Peyvandipour, N. Saberian, A. Shafi, M. Donato, and S. Draghici, "A novel computational approach for drug repurposing using systems biology," *Bioinformatics*, vol. 34, no. 16, pp. 2817–2825, 2018.
- [46] L. Liao, K. Li, K. Li, Q. Tian, and C. Yang, "Automatic density clustering with multiple kernels for high-dimension bioinformatics data," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, Kansas City, MO, USA, November 2017.
- [47] S. Saxena and S. N. Prasad, "Machine learning based sensitivity analysis for the applications in the prediction and detection of cancer disease," in *Proceedings of the IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics*, Manipal, India, August 2019.
- [48] R. Chtihakkannan, P. Kavitha, T. Mangayarkarasi, and R. Karthikeyan, "Breast cancer detection using machine learning," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 11, pp. 3123–3126, 2019.
- [49] F.-P. An, "Medical image classification algorithm based on weight initialization-sliding window fusion convolutional neural network," *Complexity*, vol. 2019, Article ID 9151670, 2019.
- [50] D. Wang, D. Ma, M. Wong, and Y. Wang, "Recent advances in surgical planning & navigation for tumor biopsy and resection," *Quantitative Imaging in Medicine and Surgery*, vol. 5, no. 5, pp. 640–648, 2015.
- [51] M. Martorelli, S. Maietta, A. Gloria, R. De Santis, E. Pei, and A. Lanzotti, "Design and analysis of 3D customized models of a human mandible," *Procedia CIRP*, vol. 49, pp. 199–202, 2016.
- [52] N. H. Cohrs, A. Petrou, M. Loeffe et al., "A soft total artificial heart—first concept evaluation on a hybrid mock circulation," *Artificial Organs*, vol. 41, no. 10, pp. 948–958, 2017.
- [53] S. H. Park, R. Su, J. Jeong et al., "3D printed polymer photodetectors," *Advanced Materials*, vol. 30, 2018.
- [54] J. Yue, P. Zhao, J. Y. Gerasimov et al., "3D-Printable Antimicrobial composite resins," *Advanced Functional Materials*, vol. 25, no. 43, pp. 6756–6767, 2015.
- [55] N. Hakimi, R. Cheng, L. Leng et al., "Handheld skin printer: in situ formation of planar biomaterials and tissues," *Lab on a Chip*, vol. 18, no. 10, 2018.
- [56] A. E. Jakus, A. L. Rutz, S. W. Jordan et al., "Hyperelastic "bone": a highly versatile, growth factor-free, osteoregenerative, scalable, and surgically friendly biomaterial," *Science Translational Medicine*, vol. 3, no. 358, 2016.
- [57] M. S. Mannoor, Z. Jiang, T. James et al., "3D printed bionic ears," *Nano Letters*, vol. 13, no. 6, pp. 2634–2639, 2013.
- [58] R. Atallah and A. Al-Mousa, "Heart disease detection using machine learning majority voting ensemble method," in *Proceedings of the 2nd International Conference on new Trends in Computing Sciences*, Amman, Jordan, October 2019.
- [59] P. S. Kohli and S. Arora, "Application of machine learning in disease prediction," in *Proceedings of the 4th International Conference on Computing Communication and Automation*, Greater Noida, India, December 2018.
- [60] C. Cao, F. Liu, H. Tan et al., "Deep learning and its applications in biomedicine," *Genomics, Proteomics & Bioinformatics*, vol. 16, no. 1, pp. 17–32, 2018.

Research Article

Comparative Evaluation of the Treatment of COVID-19 with Multicriteria Decision-Making Techniques

Figen Sarigul Yildirim ¹, **Murat Sayan** ^{2,3}, **Tamer Sanlidag** ^{3,4}, **Berna Uzun** ^{3,5},
Dilber Uzun Ozsahin ^{3,6} and **Ilker Ozsahin** ^{3,6}

¹Health Science University, Antalya Education and Research Hospital,

Department of Infectious Diseases and Clinical Microbiology, Antalya 07050, Turkey

²Faculty of Medicine, Clinical Laboratory, PCR Unit, Kocaeli University, Kocaeli, Turkey

³DESAM Institute, Near East University, 99138 Nicosia/TRNC, Mersin 10, Turkey

⁴Department of Medical Microbiology, Manisa Celal Bayar University, Manisa, Turkey

⁵Department of Mathematics, Near East University, 99138 Nicosia/TRNC, Mersin 10, Turkey

⁶Department of Biomedical Engineering, Faculty of Engineering, Near East University, 99138 Nicosia/TRNC, Mersin 10, Turkey

Correspondence should be addressed to Ilker Ozsahin; ilker.ozsahin@neu.edu.tr

Received 28 August 2020; Revised 25 December 2020; Accepted 9 January 2021; Published 23 January 2021

Academic Editor: Giovanni Improta

Copyright © 2021 Figen Sarigul Yildirim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objectives. The outbreak of coronavirus disease 2019 (COVID-19) was first reported in December 2019. Until now, many drugs and methods have been used in the treatment of the disease. However, no effective treatment option has been found and only case-based successes have been achieved so far. This study aims to evaluate COVID-19 treatment options using multicriteria decision-making (MCDM) techniques. **Methods.** In this study, we evaluated the available COVID-19 treatment options by MCDM techniques, namely, fuzzy PROMETHEE and VIKOR. These techniques are based on the evaluation and comparison of complex and multiple criteria to evaluate the most appropriate alternative. We evaluated current treatment options including favipiravir (FPV), lopinavir/ritonavir, hydroxychloroquine, interleukin-1 blocker, intravenous immunoglobulin (IVIG), and plasma exchange. The criteria used for the analysis include side effects, method of administration of the drug, cost, turnover of plasma, level of fever, age, pregnancy, and kidney function. **Results.** The results showed that plasma exchange was the most preferred alternative, followed by FPV and IVIG, while hydroxychloroquine was the least favorable one. New alternatives could be considered once they are available, and weights could be assigned based on the opinions of the decision-makers (physicians/clinicians). The treatment methods that we evaluated with MCDM methods will be beneficial for both healthcare users and to rapidly end the global pandemic. The proposed method is applicable for analyzing the alternatives to the selection problem with quantitative and qualitative data. In addition, it allows the decision-maker to define the problem simply under uncertainty. **Conclusions.** Fuzzy PROMETHEE and VIKOR techniques are applied in aiding decision-makers in choosing the right treatment technique for the management of COVID-19.

1. Introduction

Since December 2019, when coronavirus disease (COVID-19) incidents were first reported in Wuhan, China, an increasing number of cases have been reported in all countries on all continents except Antarctica. It exceeded the rate of the number of COVID-19 patients, thus prompting the World Health Organization to declare COVID-19 as an

epidemic [1]. The virus that causes COVID-19 is called the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of December 15, 2020, there were around 71,000,000 confirmed cases and 1,600,000 confirmed deaths all over the world [1].

The virus is released from respiratory secretions when an infected person speaks, sneezes, or coughs. When other people come into direct contact with these droplets, they

become infected. If those who touch the surface of the virus then touch the mouth, nose, and eyes, the infection can be transmitted to other people [2–4]. The exact incubation time is unknown. It is assumed that it is between 2 and 14 days after exposure, and most cases occur within 5 days after exposure [5, 6]. Often most infections are self-limited. It may cause more serious illness in the elderly population and those with underlying medical disease [7]. According to current statistics, the most common clinical features at the onset of illness are fever in 88%, fatigue in 38%, dry cough in 67%, myalgias in 14.9%, and dyspnea in 18.7% [8]. Pneumonia is the most common complication. Severe cases have a mortality rate of 2.3 to 5% [7].

To date, there are no proven specific treatments for patients with the new virus other than supportive care. In China, Italy, France, Spain, Turkey, and now the USA, a large number of patients have received off-label and compassionate use therapies [1]. Therefore, various options have been used to fight the virus so far. Three general methods are used, including current broad-spectrum drugs, immunoenhancement therapy, and viral-specific plasma globulin. Many drugs such as chloroquine, hydroxychloroquine, azithromycin, interferon (IFN), favipiravir (FPV), remdesivir, and lopinavir/ritonavir have been used in patients with SARS or Middle East respiratory syndrome (MERS), but the effectiveness of some drugs remains controversial [9–15]. Synthetic recombinant interferon α , intravenous immunoglobulin (IVIG) (an immunomodulator), tocilizumab, and interleukin-1 blocker are used in immunoenhancement therapy [13, 14, 16]. Convalescent plasma therapy was thought to be an effective way to alleviate the course of the disease for seriously infected patients, and successful results were obtained in patients on whom this treatment was attempted [17, 18].

These treatment methods have largely been administered in an uncontrolled manner, with the exception of a few randomized trials launched in China and recently in the USA [19]. Hence, it is not possible to make the interpretation that if the patient dies, they die from the disease, but if the patient survives, this is because of the drug given. All methods have some advantages and disadvantages. For example, chloroquine/hydroxychloroquine, azithromycin, and lopinavir-ritonavir have several negative effects, including QT prolongation, hepatitis, acute pancreatitis, neutropenia, and anaphylaxis [20, 21]. Agents that have been used less frequently in the past (e.g., remdesivir) can cause serious negative effects that have not been detected previously due to the limited number of exposed patients [22, 23]. Interleukin blockers may cause immunosuppression, increasing the risk of sepsis, bacterial pneumonia, and hepatotoxicity [24, 25].

The rapid and simultaneous combination of supportive care and randomized control studies is the only way to find effective and safe treatments for COVID-19 and other future outbreaks. In one open-label nonrandomized control study, FPV had significantly better outcomes for disease progression and viral clearance; these results should provide an important advance in creating standard treatment guidelines for combating SARS-CoV-2 infection [26].

Decision-making models are the supportive systems that will give the necessary information to decision-makers about the alternatives and their features. PROMETHEE and VIKOR techniques are commonly used analytical multicriteria decision-making techniques that rank the alternatives under the conflicting criteria successfully applied in many fields among the other techniques. VIKOR method ranks alternatives giving the compromise solution, which is the closest to the ideal solution. PROMETHEE method gives a comparison to the alternatives based on the pairwise comparison. As opposed to other methods, PROMETHEE gives more choices to the decision-maker for defining the preference function for each criterion specifically. This makes the PROMETHEE method more sensitive in ranking the alternatives.

After the fuzzy set theory has been proposed and defined by Zadeh in 1965, the hybrid models of the classical models, fuzzy-based models, have been studied by researchers in many fields. Since there is no crisp difference between many objects and cases in the real world, it has been obtained that defining and modeling the problems using fuzzy sets can create a more sensitive model to real-world problems. And fuzzy-based MCDM techniques allow the decision-makers to analyze the alternatives even with linguistic data; it is more suitable for many cases where the numerical data are not available.

Fuzzy logic has shed light to integrate human opinion into decision-making problems [27]. In 2000, Warren et al. [28] showed the applicability of the fuzzy logic in modeling the vagueness of the treatment based on the clinical guideline knowledge to support the decision-makers in clinical practices. In 2012, Consenza [29] proposed the fuzzy expert system to provide the optimal amount of the insulin unit that should be taken before the meal for the type-1 diabetes patients corresponding to the characteristic of the food. In 2017, Santini et al. [30] proposed a fuzzy-based tool in order to manage and monitor the clinical status of β -thalassemia patients. This study has given exemplary results on the online determination of iron overload while monitoring the health conditions of β -thalassemia patients. In 2019, Akram and Adeel [31] discussed the hybrid model of the hesitant m-polar fuzzy sets. They provided the application of the hesitant m-polar fuzzy TOPSIS technique to rank and select the best alternative under m-polar fuzzy set positive and negative ideal solution for a multicriteria group decision. In 2020, Akram et al. [32] proposed Dombi aggregation operators for decision-making under m-polar fuzzy information. They tested their operation validity for obtaining the best agricultural land and the best bank with its performance. They also compared their technique with the ELECTRE-I method and they found that the optimal alternative is the same by applying the ELECTRE-I method. Garg et al. [27] proposed a new fuzzy operation compared with the Yager operation for Fermatean fuzzy numbers, and they applied this technique to obtain a ranking result for the COVID-19 testing facilities. In 2020, Akram et al. [33] presented new aggregation operators such as Fermatean fuzzy Einstein weighted averaging, Fermatean fuzzy Einstein ordered weighted averaging, generalized Fermatean fuzzy

Einstein weighted averaging, and generalized Fermatean fuzzy Einstein ordered weighted averaging to cumulate the Fermatean fuzzy data in decision-making environments which has more flexible structure than the intuitionistic and Pythagorean fuzzy sets. They applied these techniques for the COVID-19 sanitizer selection problem.

Thus far, we have used a multicriteria decision-making (MCDM) technique called fuzzy PROMETHEE to compare the confusion regarding the choice of effective treatment practices using the following guidelines, to make a mutual comparison between selected treatment methods, and to determine the strongest one [34–38]. The PROMETHEE technique was developed by Brans and Vincke in 1985 [39]. Since then, it has been used successfully as an MCDM technique in many fields [40], and recently, it has been applied in the field of material selection, medicine, and healthcare [41–50], as well as it was selected as a proper COVID-19 diagnosis tool [51]. VIKOR is another commonly used MCDM technique that determines the order of alternatives under conflicting criteria based on the closeness to the ideal solution [52]. It provides the maximum group utility and the minimum individual regret for determining the compromise solution of the decision-making problem. Fuzzy logic was first presented in 1965 by Zadeh in order to define vague conditions or linguistic data mathematically [53]. A fuzzy logic-based clinical decision support system for the evaluation of renal function in posttransplant patients has been implemented successfully [54]. The analytic hierarchy process (AHP), which is another methodology based on both mathematical and psychological approaches with a similar purpose as PROMETHEE and VIKOR, is exploited to analyze and solve complex problems, in order to make the best decision. Some examples of the application of AHP in healthcare and health technologies can be found in [55–57].

The data of the COVID-19 treatment techniques contain quantitative and qualitative data of multicriteria with different importance weights which cannot be simply evaluated by the decision-makers. Therefore, we have applied two of the successfully used analytical MCDM techniques. Both of the techniques give consistent ranking results as expected. Besides, by applying the PROMETHEE method, we are able to see the positive and negative sides of each alternative.

In this study, the use of fuzzy scale based on fuzzy logic has enabled the experts to express qualitative data such as side effects, applicability, and compliance of COVID-19 treatment techniques meaningfully to be included in the model. It is also used to express the degrees of importance assigned to the criteria more easily and with the common sense of the experts.

2. Methods

SARS-CoV-2 is an enveloped, positive, single-strand RNA beta-coronavirus and is structurally similar to SARS-CoV-1 and MERS [58]. There are two overlapping hypotheses in the pathogenesis of the disease: triggered by the virus itself and the host response. There is significant confusion regarding the therapeutic approaches used in COVID-19. Currently, it is imperative to distinguish between the phase of viral

pathogenicity and the phase in which the host inflammatory response is predominant in terms of the timing of the agent to be used in the treatment. In this case, different methods have been used to treat COVID-19 at different stages of the disease. However, scientists and real-life data have demonstrated that it is more beneficial to start many treatments early [12].

2.1. Treatment Techniques. We have chosen the methods that have been most frequently used since the beginning of the COVID-19 pandemic. Chloroquine, an antimalaria drug, is the first agent used against the COVID-19 disease. It both has an immune-modulating activity and can inhibit this virus with an in vitro effect [10]. It has been proven in a clinically controlled study that chloroquine, an antimalarial drug, is effective in the treatment of patients with COVID-19 [59]. Since the side effects of hydroxychloroquine are lower than chloroquine, the use of hydroxychloroquine has been preferred in the treatment of COVID-19 [60]. In COVID-19 patients, hydroxychloroquine allows the viral nasopharyngeal carriage of SARS-CoV-2 to be cleared in just three to six days [61]. Its effect is reinforced by azithromycin, an antibacterial agent [61]. Azithromycin is used to treat or prevent certain bacterial infections, predominantly those causing middle ear infections, throat, pneumonia, typhoid, bronchitis, and sinusitis [62]. Azithromycin and hydroxychloroquine have the side effect of QT prolongation (that can cause sudden death).

Remdesivir, a terminate of viral RNA transcription, has in vitro activity against SARS-CoV-2 [63]. It was originally used for the treatment of the Ebola virus disease and Marburg virus infections. One of the side effects of remdesivir is hepatotoxicity.

FPV selectively inhibits RNA-bound RNA polymerase of RNA viruses and has been approved for new influenza therapy since 2014 [64]. It is also hepatotoxic.

Lopinavir and ritonavir are antiretroviral protease inhibitors that have been used in combination for the treatment of human immunodeficiency virus infection since 2000 [65]. This has reduced the replication by 50% in the MERS coronavirus in vitro [66]. The most common side effects are gastrointestinal problems such as nausea, vomiting, and diarrhea.

Oseltamivir is a neuraminidase enzyme inhibitor used for influenza. In fact, many patients with COVID-19 symptoms might have influenza. Therefore, it is better to give this medicine to prevent the patient from getting worse [66].

Interferon is a broad-spectrum antiviral agent that acts by interacting with toll-like receptors and inhibits viral replication [67] and anti-SARS-CoV-1 activity in vitro [68].

Tocilizumab is a recombinant humanized monoclonal antibody that acts as an IL-6 receptor antagonist and is used for the treatment of rheumatoid arthritis. Interleukin-6 was reported to be released considerably in SARS and MERS patients and might play a role in the pathogenesis of these diseases [69]. In COVID-19 patients, higher plasma levels of cytokines have also been found [69].

IVIG might be the safest immunomodulator for long-term use in patients of all ages and could help to inhibit the production of proinflammatory cytokines and increase the production of anti-inflammatory mediators [13].

While evaluating these methods, we used different criteria that were selected based on the doctors, the treatment methods, and the host. Regime cost, side effects, number of tablets, dose frequency, treatment duration, plasma stability, plasma turnover, time of suppression, practicability, intravenous form, oral form, and drug-drug interaction were chosen as the treatment method-related criteria. Age, pregnancy, glomerular filtration rate (GFR), compliance, fever, pneumonia, intensive care, organ failure, macrophage activation syndrome, and the hemophagocytic syndrome were chosen as the host-related criteria. These are symptoms and phases of diseases in COVID patients. False prescription and inefficient drug combination were chosen as the doctor-related criteria. All of these factors are important when selecting treatment methods for COVID-19 patients.

2.2. Fuzzy-Based MCDM Models. Ranking the fuzzy numbers contains the main problematic part of the decision-making problems under the fuzzy environment. This process is also the most important stage of the decision-making process since it simplifies the complexity of the problem. Comparison of the fuzzy numbers has practical applications in optimization, forecasting, decision-making, and approximate reasoning. Real-world decision problems often involve uncertain conditions of the properties of alternatives, and fuzzy numbers allow the decision-maker to be used in the analysis by taking these uncertain conditions into account. There are many types of fuzzy-based MCDM models available for different aims, and there are many studies that present different techniques for the comparison of the fuzzy numbers [70–72]. Although the centroid concept has been applied in many fields and has been known for hundreds of years, first in 1980, Yager proposed the centroid method for the comparison of the fuzzy numbers [73]. After Yager, there have been many studies that used this method for the construction of the individually defined ranking index [74–76]. Some of these researchers have used the value of x alone, while some of them used a combination of x and y values to obtain their own ranking index that depends on the centroid concept. There are also some research studies that aim to produce the most suitable or correct centroid value formula [77, 78] for the usage of the ranking index. These research studies provide valuable information for comparing the fuzzy numbers that depend on the centroid approaches. However, some studies found that the Yager index has great potential in fuzzy optimization [79]. In our study, some of the criteria, such as side effects, were defined by the experts using the linguistic fuzzy scale since no crisp values are available for those criteria. Furthermore, fuzzy data of the defined parameters have been compared using the centroid concept, defined by the Yager index, which is successfully applied and confirmed by many researchers. If we could

have used different fuzzy models most probably, we will obtain the same ranking results.

2.3. Fuzzy PROMETHEE. In real-world problems, one of the major challenges involves the collection of crisp data to analyze systems. In 1965, Zadeh laid the foundations for the idea of establishing a rule functioning and then transferring it to a machine by making use of human life experiences and various kinds of knowledge. Fuzzy logic can be defined as a decision mechanism design in its simplest form. It allows decision-makers to identify vague conditions and analyze the systems using linguistic data if it is needed [53].

MCDM is a research area that involves the analysis of various available choices in a situation or research area, which spans daily life, social sciences, engineering, medicine, and many other areas.

MCDM analyzes the alternatives (qualitatively or quantitatively) involved in a criterion that makes the alternatives a favorable or unfavorable choice for a particular application and attempts to compare this criterion based on the selected criteria against every other available option in an attempt to support the decision-maker when selecting an option with maximum advantages without negotiation.

PROMETHEE is an MCDM tool that allows a user to analyze and rank accessible alternatives according to the criteria of each alternative. It compares the available alternatives based on the selected criteria [39]. The PROMETHEE technique is a valuable and sensitive MCDM technique for reasons that include the following:

- (i) PROMETHEE can be used to handle qualitative and quantitative criteria simultaneously
- (ii) PROMETHEE deals with fuzzy relations, vagueness, and uncertainties
- (iii) PROMETHEE is easy to handle and provides the user with maximum control over the preference of the alternatives with regard to the criteria

When using PROMETHEE, only two types of information are required from the decision-maker: information regarding the importance weights of the selected criteria and the preference function to be used in comparing the alternatives' contribution with regard to each criterion.

Different preference functions (P_j) are available on PROMETHEE for the definition of different criteria. The preference function defines assigning values to the ranking of two alternatives (a and a_t) in relation to specific criteria and a preference degree of the limit between 0 and 1 [39]. The preference functions for practical purposes that can be used at the discretion of the decision-maker include the usual function, V-shape function, level function, u-shape function, linear function, and Gaussian function [39]. A complete explanation of the preference functions used, their ranking, and how to make a decision on which function best suits a scenario was presented by Brans et al. [39].

In the PROMETHEE analysis, after collecting the criteria of the alternatives, the preference function p_j (d) for each

criterion j should be defined, and the importance weight of each criterion $w_t = (w_1, w_2, \dots, w_k)$ should be determined. Normalization of the weights or equality of weights can be chosen by the decision-maker based on the application. Then, for every alternative pair $(a_t, a_{t'} \in A)$, the outranking relation $\pi(a_t, a_{t'})$ should be determined as seen in

$$\pi(a_t, a_{t'}) = \sum_{k=1}^K w_k \cdot [p_k(f_k(a_t) - f_k(a_{t'}))], \quad AXA \longrightarrow [0, 1]. \quad (1)$$

And the positive outranking flow of a_t ($\Phi^+(a_t)$) and the negative outranking flow of a_t ($\Phi^-(a_t)$) should be calculated as seen in

$$\Phi^+(a_t) = \frac{1}{n-1} \sum_{\substack{t'=1 \\ t' \neq t}}^n \pi(a_t, a_{t'}), \quad (2)$$

$$\Phi^-(a_t) = \frac{1}{n-1} \sum_{\substack{t'=1 \\ t' \neq t}}^n \pi(a_{t'}, a_t),$$

where n denotes the number of options, and each option is compared to the $n-1$ number of alternatives. The positive outranking flow is an illustration of how a particular alternative is greater than the other options. The higher the positive outranking flow of a particular alternative is, the greater the possibilities are. The negative outranking flow is an expression of how a particular alternative is less preferred compared to the other alternatives. The lower the negative outranking value is, the greater the alternatives are [1–9, 16]. PROMETHEE gives a partial preorder of the alternatives as seen in equation (3).

a_t is preferred to alternative $a_{t'}$ ($a_t Pa_{t'}$) if

$$\begin{cases} \Phi^+(a_t) \geq \Phi^+(a_{t'}) \text{ and } \Phi^-(a_t) < \Phi^-(a_{t'}), \\ \Phi^+(a_t) > \Phi^+(a_{t'}) \text{ and } \Phi^-(a_t) = \Phi^-(a_{t'}). \end{cases} \quad (3)$$

a_t is indifferent to alternative $a_{t'}$ ($a_t Ia_{t'}$) if

$$\Phi^+(a_t) = \Phi^+(a_{t'}) \text{ and } \Phi^-(a_t) = \Phi^-(a_{t'}). \quad (4)$$

a_t is incomparable to $a_{t'}$ ($a_t Ra_{t'}$) if

$$\begin{cases} \Phi^+(a_t) > \Phi^+(a_{t'}) \text{ and } \Phi^-(a_t) > \Phi^-(a_{t'}), \\ \Phi^+(a_t) < \Phi^+(a_{t'}) \text{ and } \Phi^-(a_t) < \Phi^-(a_{t'}). \end{cases} \quad (5)$$

Using PROMETHEE II, the net flow of an alternative a_t ($\Phi^{\text{net}}(a_t)$) can be calculated with

$$\Phi^{\text{net}}(a_t) = \Phi^+(a_t) - \Phi^-(a_t). \quad (6)$$

And the net ranking results of the alternatives can be found by

$$(a_t Pa_{t'}) \text{ if } \Phi^{\text{net}}(a_t) > \Phi^{\text{net}}(a_{t'}), \quad (7)$$

$$(a_t Ia_{t'}) \text{ if } \Phi^{\text{net}}(a_t) = \Phi^{\text{net}}(a_{t'}). \quad (8)$$

The better alternative is the one with the higher $\Phi^{\text{net}}(a_t)$ value (32).

In this study, we proposed the use of the fuzzy-based PROMETHEE technique for the evaluation of the available treatment options for COVID-19. The selected treatment options are favipiravir (FPV), remdesivir, “lopinavir/ritonavir,” “hydroxychloroquine,” “oseltamivir + hydroxychloroquine,” “hydroxychloroquine + azithromycin,” interleukin-1 blocker, tocilizumab, “interferon,” intravenous immunoglobulin (IVIG), and plasma exchange. The selected criteria of the alternatives and the importance weight with fuzzy scale can be seen in Table 1. The data and the weights of the criteria have been collected by the experts.

The Yager index was used for the defuzzification of the fuzzy scale. Lastly, the PROMETHEE-Gaia decision lab program with Gaussian preference functions has been used for the evaluation.

2.4. Fuzzy VIKOR. The basis of the VIKOR (Multicriteria Optimization and Compromise Solution) method is based on the determination of a compromise solution in the light of alternatives and within the scope of the evaluation criteria. This compromise solution has been determined as the closest solution to the ideal solution [80, 81]. With the expression of a compromise solution, it is understood that by creating a multicriteria ranking index for alternatives, the closest decision is made to the ideal solution under certain conditions. Under the assumption that each alternative is evaluated on the basis of decision-making criteria, consensus ranking is achieved by comparing the values of proximity to the ideal alternative [82]. This technique is also based on obtaining the ranking results of alternatives with maximum group benefit and therefore minimum regret of individuals.

After constructing the decision matrix of the MCM problem with specifying the importance weights of each criterion, the steps of the VIKOR method can be summarized as follows.

Step 1 (determination of the best (f_i^*) and the worst (f_i^-) values of each criterion). The best values should be determined as the maximum point of beneficial criteria and the minimum point of the criteria that cause cost. If the criterion- i is the beneficial criterion, then f_i^* and f_i^- can be obtained by using the following formulas:

$$f_i^* = \max_j(f_{ij}), \quad (9)$$

$$f_i^- = \min_j(f_{ij}). \quad (10)$$

f_{ij} denotes the value of the j -th criterion of the i -th alternative.

Step 2 (obtaining the Utility (S_i) and Regret (R_i) values for each alternative). The utility (S_i) and regret (R_i) values of the alternative- i can be calculated by using the following formulas:

TABLE 1: Criteria of the COVID-19 treatment options and their importance weights with fuzzy linguistic scale.

Linguistic scale for evaluation	Triangular fuzzy scale	Importance ratings of criteria
Very high (VH)	(0.75, 1, 1)	Side effects, regime cost, number of tablets, dose frequency, treatment duration, plasma stability, plasma turnover, time of suppression, practicability, drug-drug interaction, compliance, fever, pneumonia, intensive care, organ failure, macrophage activation syndrome, hemophagocytic syndrome
Important (H)	(0.50, 0.75, 1)	Age, pregnancy, GFR
Medium (M)		Intravenous form, oral form, false prescription
Low (L)		Inefficient drug combination
Very low (VL)		-

$$S_i = \sum_{i=1}^n w_i \left[\frac{f_i^* - f(ij)}{f_i^* - f_i^-} \right], \quad (11)$$

$$R_j = \max_j \left(w_i \left[\frac{f_i^* - f(ij)}{f_i^* - f_i^-} \right] \right), \quad (12)$$

where w_i denotes the importance weights of the criterion- j , which represents the relative importance degrees.

Step 3 (computing the value of Q_j and ranking the alternatives accordingly). Q_j values can be obtained based on the relation given as

$$Q_j = \nu \left[\frac{S_j - \min(S_j)}{\max(S_j) - \min(S_j)} \right] + (1 - \nu) \left[\frac{R_j - \min(R_j)}{\max(R_j) - \min(R_j)} \right], \quad (13)$$

where $\nu \in [0, 1]$ and represents the weight/level of the strategy that indicates the maximum group utility and $(1 - \nu)$ represents the weight/level of the individual regret. The value of ν is most commonly used as 0.5, so as in this study.

When ν value (>0.5) is chosen high, it is stated that the majority tend to show a positive attitude toward Q_j index. When the ν value (<0.5) is chosen less, it means that the majority shows a negative attitude. In general, it is assumed that the evaluation expert groups (positive and negative) have a conciliatory attitude by selecting the ν value = 0.5 [83].

The alternative with the smallest Q_j value is indicated as the best option within the group of alternatives. And, the ranking can be obtained accordingly.

However, in order to consider the alternative with a minimum value of Q_j as the best alternative with an acceptable advantage, it must meet the following conditions:

Condition 1 (acceptable advantage): the acceptable advantage is the existence of a significant difference between the best and the closest options, which should be calculated as

$$Q(A'') - Q(A') \geq DQ. \quad (14)$$

A'' denotes the best alternative with the minimum Q value and A' denotes the second-best alternative with the second minimum Q_j .

$$DQ = \frac{1}{(m - 1)}, \quad (15)$$

where m indicates the number of the alternatives.

Condition 2 (acceptable stability): A' should have the minimum (best) value/s of the R_j and/or S_j between all of the alternatives.

In VIKOR, if the best alternative with the min (Q_j) does not satisfy one of the given conditions, then the compromise solutions set can be proposed as follows:

- (i) If only condition 2 is not satisfied, one has A' and A'
- (ii) If the first condition is not satisfied, one has A', A', \dots, A^M where M represents the maximum decision points that meet the condition

$$Q(A^M) - Q(A') < DQ. \quad (16)$$

3. Results

Table 2 shows the complete ranking results of the COVID-19 treatment options with the positive, negative, and net flow ranking of the alternatives. According to the table, plasma exchange and FVP outrank the other treatment options with net flows of 0.0268 and 0.0265, respectively, followed by IVIG. The least ranked COVID-treatment option is hydroxychloroquine with a net flow of -0.0502 . Similarly, Figure 1 shows the strong and weak points of the COVID-19 treatment options.

These results of the VIKOR technique validate the ranking results of the PROMETHEE technique. Using the VIKOR method, we obtained almost the same net ranking for the selected COVID-19 treatment alternatives as seen in Table 3.

4. Discussion

In our study, we compared the treatment applications used in COVID-19 treatment between December 2019 and March 2020 with MCDM methods. Clinical studies of the treatment methods used are still ongoing. In our study, the plasma exchange method was found to be the best method among the treatment options, similar to clinical applications. It has

TABLE 2: Complete ranking of COVID-19 treatment options with PROMETHEE.

Complete ranking	Options (a_i)	Net flow $\Phi_i^{net}(a_i)$	Positive flow $\Phi_i^+(a_i)$	Negative flow $\Phi_i^-(a_i)$
1	Plasma exchange	0.0268	0.0364	0.0095
2	Favipravir	0.0265	0.0361	0.0096
3	Intravenous immunoglobulin (IVIG)	0.0199	0.0318	0.0120
4	Interleukin-1 blocker	0.0189	0.0335	0.0135
5	Tocilizumab	0.0184	0.0342	0.0158
6	Remdesivir	0.0177	0.0340	0.0162
7	Interferon	0.0139	0.0283	0.0144
8	Oseltamivir + hydroxychloroquine + azithromycin	-0.0188	0.0165	0.0352
9	Lopinavir/ritonavir	-0.0348	0.0239	0.0588
10	Oseltamivir + hydroxychloroquine	-0.0384	0.0103	0.0487
11	Hydroxychloroquine	-0.0502	0.0108	0.0609

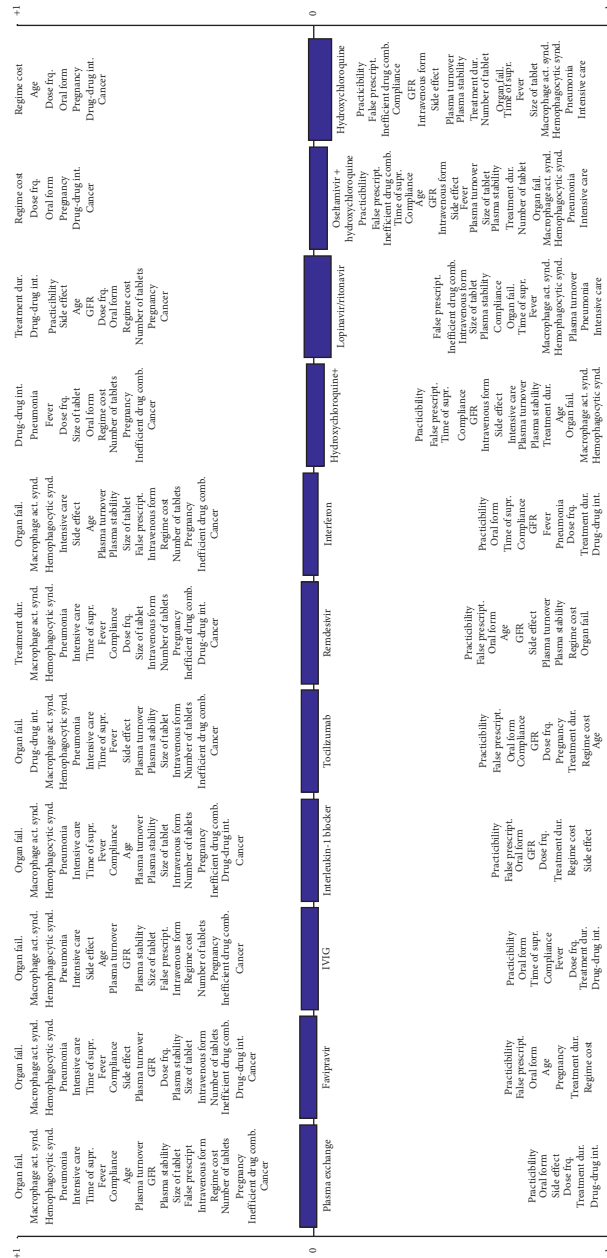


FIGURE 1: Positive and negative aspects of each COVID-19 treatment option. The higher the criterion stands in the graph in the positive side, the higher it contributes to the positive side of the technique. Similarly, the lower the criterion stands in the graph on the negative side, the higher it contributes to the negative side of the technique.

TABLE 3: Complete ranking of COVID-19 treatment options with VIKOR.

Complete ranking	Alternatives	Qj	Sj	Rj
1	Plasma exchange	0	5.1	0.92
2	Favipravir	0.013387978	5.345	0.92
3	Intravenous immunoglobulin	0.050273224	6.02	0.92
4	Interleukin-1 blocker	0.093442623	6.81	0.92
5	Remdesivir	0.113934426	7.185	0.92
6	Interferon	0.116393443	7.23	0.92
7	Tocilizumab	0.125136612	7.39	0.92
8	Oseltamivir + hydroxychloroquine + azithromycin	0.310382514	10.78	0.92
9	Lopinavir/ritonavir	0.337978142	11.285	0.92
10	Oseltamivir + hydroxychloroquine	0.449726776	13.33	0.92
11	Hydroxychloroquine	0.5	14.25	0.92

been suggested that plasma from cured patients should be used for treatment [84]. Indeed, healing patients often have a high level of antibodies to the pathogen present in their blood. Antibodies are immunoglobulin produced by B lymphocytes to fight pathogens and other foreign bodies, to recognize and neutralize unique molecules in pathogens [85]. Based on this, patients who recovered from COVID-19 recovered and were injected plasma into serious patients after collection and processing, and within 24 hours, there was a decrease in inflammation and viral loads as well as an improvement in oxygen saturation in the blood [86].

FVP was found to be the second-best method of treatment for COVID-19 patients. The patients who were treated with FVP showed faster viral clearance, shorter fever period, and improvement in radiological findings in Wuhan [87]. The third best one was found to be IVIG which may also play a role in the control of the immune system, where there is a high level of inflammation. Improvement in the poor prognosis stage of the disease is poor, and IVIG can be quickly recognized and applied for this treatment [13].

When we compared the methods related to the criteria, the others were interleukin-1 blocker, remdesivir, interferon, tocilizumab,

oseltamivir + hydroxychloroquine + azithromycin, lopinavir/ritonavir, oseltamivir + hydroxychloroquine, and hydroxychloroquine. These criteria are very important, not only for patients but also for treatment practitioners.

The spread of COVID-19 is continuing at a rapid pace. It is very important to discover effective treatment or prophylactic agents among measures such as staying at home, hand hygiene, wearing masks, and so on to stop the pandemic. Health professionals and the global scientific community are waiting for new evidence to emerge soon to manage COVID-19. Until this evidence emerges, it is necessary to continue using the treatment methods that have shown effectiveness. The treatment methods that we evaluated with MCDM methods will be beneficial for both healthcare users and to rapidly end the global pandemic.

In this study, an average patient was considered to show the method's applicability, so the patient profile such as gender or disease stage was not included in the analysis. However, this study can be extended to include all possible factors since fuzzy PROMETHEE and VIKOR are able to support a large number of inputs. Another limitation of this

study is the fact that treatment selection might be different for patients in the acute phase than those in the stable phase which was not considered in this study.

The proposed method is applicable for analyzing the alternatives to the selection problem with quantitative and qualitative data. In addition, it allows the decision-maker to define the problem simply under uncertainty. One of the limitations of this technique is that there is no method for determining the weight of the criteria. Therefore, expert opinion is of great importance in establishing criteria weights for this model to give accurate results in practical applications.

5. Conclusion

In this study, we analyzed different treatment options for COVID-19 treatment using fuzzy PROMETHEE and VIKOR methods. Overall, there is no globally approved specific antiviral drug available for COVID-19. All drug options come from the experience of treating SARS, MERS, or other new influenza viruses. Active symptomatic support is the key to treatment. The above medicines will help, and their efficacy needs further confirmation. New alternatives and criteria could be considered once they are available in the future, and weights could be assigned based on the opinions of the decision-makers (physicians/clinicians). We showed the applicability of the MCDM techniques in informing decision-makers in terms of choosing the right treatment technique for the management of COVID-19.

Data Availability

All data generated or analyzed during this study are included within this published article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

F. S. Y. and B. U. wrote the first draft of the manuscript. B. U., D. U. O., and I. O. analyzed the data. All the authors proposed the main ideas for the research, provided

interpretations, and critically commented, revised, and approved the final manuscript.

References

- [1] World Health Organization, "Situation updates on March 26, 2020," 2020, <https://covid19.who.int/>.
- [2] C. Rothe, M. Schunk, P. Sothmann et al., "Transmission of 2019-nCoV infection from an asymptomatic contact in Germany," *New England Journal of Medicine*, vol. 382, no. 10, pp. 970–971, 2020.
- [3] K. Kupferschmidt, "Study claiming new coronavirus can be transmitted by people without symptoms was flawed science," *The Journal of Infectious Diseases*, vol. 3, 2020.
- [4] Y. Bai, L. Yao, T. Wei et al., "Presumed asymptomatic carrier transmission of COVID-19," *JAMA*, vol. 323, no. 14, p. 1406, 2020.
- [5] Q. Li, X. Guan, P. Wu et al., "Early transmission dynamics in wuhan, China, of novel coronavirus-infected pneumonia," *New England Journal of Medicine*, vol. 382, no. 13, pp. 1199–1207, 2020.
- [6] J. F.-W. Chan, S. Yuan, K.-H. Kok et al., "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514–523, 2020.
- [7] Y. Yi, P. N. P. Lagniton, S. Ye, E. Li, and R.-H. Xu, "COVID-19: what has been learned and to be learned about the novel coronavirus disease," *International Journal of Biological Sciences*, vol. 16, no. 10, pp. 1753–1766, 2020.
- [8] N. Chen, M. Zhou, X. Dong et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *The Lancet*, vol. 395, no. 10223, pp. 507–513, 2020.
- [9] M. L. Holshue, C. DeBolt, S. Lindquist et al., "First case of 2019 novel coronavirus in the United States," *New England Journal of Medicine*, vol. 382, no. 10, pp. 929–936, 2020.
- [10] M. Wang, R. Cao, L. Zhang et al., "Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro," *Cell Research*, vol. 30, no. 3, pp. 269–271, 2020.
- [11] J. Gao, Z. Tian, and X. Yang, "Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies," *BioScience Trends*, vol. 14, no. 1, pp. 72–73, 2020.
- [12] H. Lu, "Drug treatment options for the 2019-new coronavirus (2019-nCoV)," *BioScience Trends*, vol. 14, no. 1, pp. 69–71, 2020.
- [13] L. Gilardin, J. Bayry, and S. V. Kaveri, "Intravenous immunoglobulin as clinical immune-modulating therapy," *Canadian Medical Association Journal*, vol. 187, no. 4, pp. 257–264, 2015.
- [14] V. Kumar, Y.-S. Jung, and P.-H. Liang, "Anti-SARS coronavirus agents: a patent review (2008-present)," *Expert Opinion on Therapeutic Patents*, vol. 23, no. 10, pp. 1337–1348, 2013.
- [15] S. Mustafa, H. Balkhy, and M. N. Gabere, "Current treatment options and the role of peptides as potential therapeutic components for Middle East Respiratory Syndrome (MERS): a review," *Journal of Infection and Public Health*, vol. 11, no. 1, pp. 9–17, 2018.
- [16] Actemra (tocilizumab), "Prescribing information. Genentech," 2019, https://www.actemrahcp.com/?_ga=2.20137041460.509331555.1584929819-505112783.
- [17] J. Mair-Jenkins, M. Saavedra-Campos, J. K. Baillie et al., "The effectiveness of convalescent plasma and hyperimmune immunoglobulin for the treatment of severe acute respiratory infections of viral etiology: a systematic review and exploratory meta-analysis," *Journal of Infectious Diseases*, vol. 211, no. 1, pp. 80–90, 2015.
- [18] L. Zhang and Y. Liu, "Potential interventions for novel coronavirus in China: a systematic review," *Journal of Medical Virology*, vol. 92, no. 5, pp. 479–490, 2020.
- [19] Adaptive COVID-19 treatment trial, "ClinicalTrials.gov identifier: nct04280705," 2020, <https://clinicaltrials.gov/ct2/show/NCT04280705?%20Term=remdesivir&cond=covid-19&draw=2&rank=5>.
- [20] B. E. Young, S. W. X. Ong, S. Kalimuddin et al., "Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore novel coronavirus outbreak research team. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore," *Jama*, vol. 323, no. 15, p. 1488, 2020.
- [21] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China," *Jama*, vol. 323, no. 13, p. 1239, 2020.
- [22] J. Reina, "The antiviral hope against SARS-CoV-2," *Revista Española de Quimioterapia*, vol. 33, no. 3, pp. 176–179, 2020.
- [23] K. Kupferschmidt and J. Cohen, "Race to find COVID-19 treatments accelerates," *Science*, vol. 367, no. 6485, pp. 1412–1413, 2020.
- [24] A. C. Kalil and J. Sun, "Low-dose steroids for septic shock and severe sepsis: the use of Bayesian statistics to resolve clinical trial controversies," *Intensive Care Medicine*, vol. 37, no. 3, pp. 420–429, 2011.
- [25] A. tocilizumab, "Prescribing information. Genentech," 2019, https://www.actemrahcp.com/?_ga=2.20137041460.509331555.1584929819-505112783.%201584929819.
- [26] Q. Cai, M. Yang, D. Liu et al., *Experimental Treatment with Favipiravir for COVID-19: An Open-Label Control Study Engineering*, National Library of Medicine, Beijing, China.
- [27] H. Garg, G. Shahzadi, and M. Akram, "Decision-making analysis based on Fermatean fuzzy Yager aggregation operators with application in COVID-19 testing facility," *Mathematical Problems in Engineering*, vol. 2020, Article ID 7279027, 16 pages, 2020.
- [28] J. Warren, G. Beliakov, and B. van der Zwaag, "Fuzzy logic in clinical practice decision support systems," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Maui, HI, USA, 2000.
- [29] B. Cosenza, "Off-line control of the postprandial glycemia in type 1 diabetes patients by a fuzzy logic decision support," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10693–10699, 2012.
- [30] S. Santini, "Using fuzzy logic for improving clinical daily-care of β -thalassemia patients," in *Proceedings of the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6, Naples, Italy, 2017.
- [31] M. Akram and A. Adeel, "Novel TOPSIS method for group decision-making based on hesitant m-polar fuzzy model," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 6, pp. 8077–8096, 2019.
- [32] M. Akram, N. Yaqoob, G. Ali, and W. Chammam, "Extensions of Dombi aggregation operators for decision making under m-polar fuzzy information," *Journal of Mathematics*, vol. 2020, Article ID 4739567, 2020.
- [33] M. Akram, G. Shahzadi, and A. A. H. Ahmadini, "Decision-making framework for an effective sanitizer to reduce

- COVID-19 under fermatean fuzzy environment,” *Journal of Mathematics*, vol. 2020, Article ID 3263407, 19 pages, 2020.
- [34] B. Uzun, F. Sarigül Yıldırım, M. Sayan, T. Şanlıdağ, and D. Uzun Ozsahin, *The Use Of Fuzzy Promethee Technique In Antiretroviral Combination Decision In Pediatric Hiv Treatments*, IEEE Xplore, New York, NY, USA, 2019.
- [35] M. Sayan, N. Sultanoglu, B. Uzun, F. S. Yıldırım, T. Şanlıdağ, and D. U. Ozsahin, *Determination of Post-Exposure Prophylaxis Regimen in the Prevention of Potential Pediatric HIV-1 Infection by the Multi-Criteria Decision-Making Theory*, IEEE Xplore, New York, NY, USA, 2019.
- [36] N. Sultanoglu, B. Uzun, F. S. Yıldırım, M. Sayan, T. Şanlıdağ, and D. U. Ozsahin, *Selection of the Most Appropriate Antiretroviral Medication in Determined Aged Groups (≥ 3 Years) of HIV-1 Infected Children*, IEEE Xplore, New York, NY, USA, 2019.
- [37] M. Sayan, T. Sanlidag, N. Sultanoglu, B. Uzun, F. S. Yildirim, and D. U. Ozsahin, “Evaluating the efficacy of adult HIV post exposure prophylaxis regimens in relation to transmission risk factors by multi criteria decision method,” in *Advances in Intelligent Systems and Computing*, R. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. Babanli, and F. Sadikoglu, Eds., vol. 1095, Springer, Cham, Switzerland, 2020.
- [38] M. Sayan, D. Uzun Ozsahin, T. Sanlidag, N. Sultanoglu, F. Sarigul Yildirim, and B. Uzun, “Efficacy evaluation of antiretroviral drug combinations for HIV-1 treatment by using the fuzzy PROMETHEE,” in *Advances in Intelligent Systems and Computing*, R. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. Babanli, and F. Sadikoglu, Eds., vol. 1095, Springer, Cham, Switzerland, 2020.
- [39] J. P. Brans and P. Vincke, “A preference ranking Organisation method: the PROMETHEE method for MCDM,” *Management Science*, vol. 31, no. 6, 1985.
- [40] A. Asemi, M. S. Baba, A. Asemi, R. Abdullah, and N. Idris, “Fuzzy multi criteria decision making applications: a review study,” in *Proceedings of the 3rd International Conference on Computer Engineering & Mathematical Sciences (ICCEMS 2014)*, Langkawi, Malaysia, 2014.
- [41] I. Ozsahin, S. Abebe, and G. Mok, “A multi-criteria decision-making approach for schizophrenia treatment techniques,” *Archives of Psychiatry and Psychotherapy*, vol. 22, no. 2, pp. 52–61, 2020.
- [42] I. Ozsahin, “Identifying a personalized anesthetic with fuzzy PROMETHEE,” *Healthcare Informatics Research*, vol. 26, no. 3, pp. 201–211, 2020.
- [43] I. Ozsahin, D. Uzun Ozsahin, M. Maisaini, and P. Mok, “Fuzzy PROMETHEE analysis of leukemia treatment techniques,” 2019, <https://www.wcrj.net/article/1315>.
- [44] I. Ozsahin, D. Uzun Ozsahin, K. Nyakuwanikwa, and T. Wallace Simbanegav, *Fuzzy PROMETHEE for Ranking Pancreatic Cancer Treatment Techniques*, IEEE Xplore, New York, NY, USA, 2019.
- [45] M. Maisaini, B. Uzun, I. Ozsahin, and D. Uzun, “Evaluating lung cancer treatment techniques using fuzzy PROMETHEE approach,” in *Proceedings of the 13th International Conference On Theory And Application Of Fuzzy Systems And Soft Computing — ICAFS-2018*, pp. 209–215, Warsaw, Poland, 2018.
- [46] D. Uzun Ozsahin and I. Ozsahin, “A fuzzy PROMETHEE approach for breast cancer treatment techniques,” *Health Sciences*, vol. 7, no. 5, pp. 29–32, 2018.
- [47] D. Uzun, B. Uzun, M. Sani, and I. Ozsahin, “Evaluating X-ray based medical imaging devices with fuzzy preference ranking organization method for enrichment evaluations,” *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 3, 2018.
- [48] D. Ozsahin, N. Isa, B. Uzun, and I. Ozsahin, “Effective analysis of image reconstruction algorithms in nuclear medicine using fuzzy PROMETHEE,” in *Proceedings of the 2018 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1–5, Dubai, United Arab Emirates, 2019.
- [49] I. Ozsahin, T. Sharif, D. U. Ozsahin, and B. Uzun, “Evaluation of solid-state detectors in medical imaging with fuzzy PROMETHEE,” *Journal of Instrumentation*, vol. 14, no. 1, Article ID C01019, 2019.
- [50] M. Taiwo Mubarak, I. Ozsahin, and D. Uzun Ozsahin, *Evaluation of Sterilization Methods for Medical Devices*, pp. 1–4, IEEE, New York, NY, USA, 2019.
- [51] M. Sayan, T. Sanlidag, U. Berna, and I. Ozsahin, “Capacity evaluation of diagnostic tests for COVID-19 using multi-criteria decision-making techniques,” *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 1560250, 8 pages, 2020.
- [52] N. Zhang and G. Wei, “Extension of VIKOR method for decision making problem based on hesitant fuzzy set,” *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 4938–4947, 2013.
- [53] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [54] G. Improta, V. Mazzella, D. Vecchione, S. Santini, and M. Triassi, “Fuzzy logic-based clinical decision support system for the evaluation of renal function in post-Transplant Patients,” *Journal of Evaluation in Clinical Practice*, vol. 26, no. 4, pp. 1224–1234, 2020.
- [55] G. mprota, G. Converso, T. Murino, G. Mosè, P. Antonietta, and R. Maria, “Analytic hierarchy process (AHP) in dynamic configuration as a tool for health technology assessment (HTA): the case of biosensing optoelectronics in oncology,” *International Journal of Information Technology & Decision Making*, vol. 18, pp. 1533–1550, 2019.
- [56] G. Improta, A. Perrone, M. A. Russo, and T. Maria, “Health technology assessment (HTA) of optoelectronic biosensors for oncology by analytic hierarchy process (AHP) and Likert scale,” *BMC Medical Research Methodology*, vol. 19, p. 140, 2019.
- [57] G. Improta, M. A. Russo, M. Triassi, G. Converso, T. Murino, and L. C. Santillo, “Use of the AHP methodology in system dynamics: modelling and simulation for health technology assessments to determine the correct prosthesis choice for hernia diseases,” *Mathematical Biosciences*, vol. 299, pp. 19–27, 2018.
- [58] R. Lu, X. Zhao, J. Li et al., “Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding,” *The Lancet*, vol. 395, no. 10224, pp. 565–574, 2020.
- [59] Multicenter collaboration group of Department of Science and Technology of Guangdong Province, “Expert consensus on chloroquine phosphate for the treatment of novel coronavirus pneumonia,” *Zhonghua*, vol. 43, no. 3, pp. 185–188, 2020.
- [60] M. F. Marmor, U. Kellner, T. Y. Y. Lai, R. B. Melles, and W. F. Mieler, “Recommendations on screening for chloroquine and hydroxychloroquine retinopathy (2016 revision),” *Ophthalmology*, vol. 123, no. 6, pp. 1386–1394, 2016.
- [61] P. Gautret, J.-C. Lagier, P. Parola et al., “Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial,” *International Journal of Antimicrobial Agents*, vol. 56, no. 1, p. 105949, 2020.

- [62] A. H. H. Bakheit, B. M. H. Al-Hadiya, and A. A. Abd-Elgalil, "Azithromycin," *Profiles of Drug Substances, Excipients and Related Methodology*, vol. 39, pp. 1–40, 2014.
- [63] Y.-C. Cao, Q.-X. Deng, and S.-X. Dai, "Remdesivir for severe acute respiratory syndrome coronavirus 2 causing COVID-19: an evaluation of the evidence," *Travel Medicine and Infectious Disease*, vol. 35, Article ID 101647, 2020.
- [64] Y. Furuta, K. Takahashi, M. Kuno-Maekawa et al., "Mechanism of action of T-705 against influenza virus," *Antimicrobial Agents and Chemotherapy*, vol. 49, no. 3, pp. 981–986, 2005.
- [65] R. S. Cvetkovic and K. L. Goa, "Lopinavir/ritonavir," *Drugs*, vol. 63, no. 8, pp. 769–802, 2003.
- [66] H. Momattin, A. Y. Al-Ali, and J. A. Al-Tawfiq, "A systematic review of therapeutic agents for the treatment of the Middle East respiratory syndrome coronavirus (MERS-CoV)," *Travel Medicine and Infectious Disease*, vol. 30, pp. 9–18, 2019.
- [67] S. Uematsu and S. Akira, "Toll-like receptors and type I interferons," *Journal of Biological Chemistry*, vol. 282, no. 21, pp. 15319–15323, 2007.
- [68] U. Ströher, A. DiCaro, Y. Li et al., "Severe acute respiratory syndrome-related coronavirus is inhibited by interferon- α ," *The Journal of Infectious Diseases*, vol. 189, no. 7, pp. 1164–1167, 2004.
- [69] C. Huang, Y. Wang, X. Li et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
- [70] R. Yager, "A procedure for ordering fuzzy subsets of the unit interval," *Information Science*, vol. 24, no. 1, pp. 143–161, 1981.
- [71] X. Wang and E. E. Kerre, "Reasonable properties for the ordering of fuzzy quantities (I)," *Fuzzy Sets System*, vol. 118, no. 1, pp. 375–387, 2001.
- [72] K.-P. Chiao, "A new ranking approach for general interval type-2 fuzzy sets using extended alpha cuts representation," in *Proceedings of the 22nd International Conference*, vol. 1, pp. 594–597, Belgrade, Serbia, 2015.
- [73] R. R. Yager, "On a general class of fuzzy connectives," *Fuzzy Sets and Systems*, vol. 4, no. 6, pp. 235–242, 1980.
- [74] S.-J. Chen and S.-M. Chen, "Fuzzy risk analysis based on the ranking of generalized trapezoidal fuzzy numbers," *Applied Intelligence*, vol. 26, no. 1, pp. 1–11, 2007.
- [75] C. Liang, J. Wu, and J. Zhang, *Ranking Indices and Rules for Fuzzy Numbers Based on Gravity Center Point*, Intelligent Control and Automation, Dalian, China, 2006.
- [76] Y.-J. Wang and H.-S. Lee, "The revised method of ranking fuzzy numbers with an area between the centroid and original points," *Computers & Mathematics with Applications*, vol. 55, no. 9, pp. 2033–2042, 2008.
- [77] B. S. Shieh, "An approach to centroids of fuzzy numbers," *International Journal of Fuzzy Systems*, vol. 9, no. 1, pp. 51–54, 2007.
- [78] Y.-M. Wang, J.-B. Yang, D.-L. Xu, and K.-S. Chin, "On the centroids of fuzzy numbers," *Fuzzy Sets and Systems*, vol. 157, no. 7, pp. 919–926, 2006.
- [79] J. C. Figueroa-Garcia, Y. Chalco-Cano, and H. Roman-Flores, "Yager index and ranking for interval type-2 fuzzy numbers," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 2709–2718, 2018.
- [80] S. Opricovic and G.-H. Tzeng, "Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS," *European Journal of Operational Research*, vol. 156, no. 2, pp. 445–455, 2004.
- [81] L. Y. Chen and T.-C. Wang, "Optimizing partners' choice in IS/IT outsourcing projects: the strategic decision of fuzzy VIKOR," *International Journal of Production Economics*, vol. 120, no. 1, pp. 233–242, 2009.
- [82] S. Opricovic and G.-H. Tzeng, "Extended VIKOR method in comparison with outranking methods," *European Journal of Operational Research*, vol. 178, no. 2, pp. 514–529, 2007.
- [83] J. Wei and X. Lin, "The multiple attribute decision-making VIKOR method and its application," in *Proceedings of the Wireless Communications, Networking and Mobile Computing, WiCOM'08. 4th International Conference*, pp. 1–4, IEEE, Dalian, China, 2008.
- [84] [Internet] BLOOMBERG, "China seeks plasma from recovered patients as virus treatment," <https://time.com/5784286/covid-19-chinaplasm-%20treatment/>.
- [85] T. R. Kreil and M. R. Farcet, "Immunoglobulins and virus antibody titers: of past needs, current requirements, and future options," *Transfusion*, vol. 58, no. 3, pp. 3090–3095, 2018.
- [86] C. Shen, Z. Wang, F. Zhao et al., "Treatment of 5 critically ill patients with COVID-19 with convalescent plasma," *JAMA*, vol. 323, no. 16, p. 1582, 2020.
- [87] J. Bryner, "Flu drug used in Japan shows promise in treating COVID-19," <http://www.Livescience.com>.

Research Article

Exploiting Multiple Optimizers with Transfer Learning Techniques for the Identification of COVID-19 Patients

Zeming Fan ¹, Mudasir Jamil ¹, Muhammad Tariq Sadiq ¹, Xiwei Huang ²,
and Xiaojun Yu ¹

¹School of Automation, Northwestern Polytechnical University, Xi'an 710129, China

²Ministry of Education, Key Laboratory of RF Circuits and Systems, Hangzhou Dianzi University, Hangzhou 310018, China

Correspondence should be addressed to Xiaojun Yu; e070035@e.ntu.edu.sg

Received 28 June 2020; Revised 8 October 2020; Accepted 13 November 2020; Published 24 November 2020

Academic Editor: Antonio Gloria

Copyright © 2020 Zeming Fan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the rapid spread of COVID-19 and its induced death worldwide, it is imperative to develop a reliable tool for the early detection of this disease. Chest X-ray is currently accepted to be one of the reliable means for such a detection purpose. However, most of the available methods utilize large training data, and there is a need for improvement in the detection accuracy due to the limited boundary segment of the acquired images for symptom identifications. In this study, a robust and efficient method based on transfer learning techniques is proposed to identify normal and COVID-19 patients by employing small training data. Transfer learning builds accurate models in a timesaving way. First, data augmentation was performed to help the network for memorization of image details. Next, five state-of-the-art transfer learning models, AlexNet, MobileNetv2, ShuffleNet, SqueezeNet, and Xception, with three optimizers, Adam, SGDM, and RMSProp, were implemented at various learning rates, $1e-4$, $2e-4$, $3e-4$, and $4e-4$, to reduce the probability of overfitting. All the experiments were performed on publicly available datasets with several analytical measurements attained after execution with a 10-fold cross-validation method. The results suggest that MobileNetv2 with Adam optimizer at a learning rate of $3e-4$ provides an average accuracy, recall, precision, and F -score of 97%, 96.5%, 97.5%, and 97%, respectively, which are higher than those of all other combinations. The proposed method is competitive with the available literature, demonstrating that it could be used for the early detection of COVID-19 patients.

1. Introduction

Novel coronavirus, also known as COVID-19, emerged from the city of Wuhan, China, in late 2019. A disease that appeared to be like regular flu at first has now been officially declared as pandemic and has affected more than five million people so far around the world [1, 2]. Researchers believe that COVID-19 is a type of severe acute respiratory syndrome coronavirus 2 or SARS-CoV-2 [3]. Globally, coronavirus cases have crossed five million in numbers, with the death toll surpassing 320,000 [4]. The World Health Organization (WHO) declared a worldwide health emergency on Jan 30, 2020, to point out the alarming situation. Since then, most of the countries are under lockdown, with virus cases still inclining at a rapid rate.

The health system of many developed countries reached a point near collapse due to the pandemic [5]. Even with the

latest medical facility available at hand, the only option that seems to be working against this disease is social distancing. Latest stats [6] have shown that China has efficiently defeated the virus through strict precautions and social distancing; however, the United States of America, Spain, Italy, France, and many other countries took a devastating hit from the virus. Doctors are using chest scans to diagnose the symptoms of COVID-19 rather than waiting for blood results [7]. Patients suffering from this disease have shown common signs like open holes in the lungs [8]. This research is being used to classify the patients and distinguish them from healthy people. Computed Tomography (CT) and chest X-ray provide huge assistance to doctors in diagnosing an infectious disease [9–11]. Obtaining scanned images through X-ray is faster, simpler, and economical as compared to CT scans [5]. In China, strict precautions were taken at the start to overcome the spread of this disease.

Patients who showed mild level symptoms were quarantined and tested multiple times days apart to ensure the safety of other people [12].

Due to the severance of COVID-19, many researchers are proposing different methods to identify the infection through early symptoms and take the required measurements [13]. Even with so many available methods for the detection of the COVID-19 symptoms, there can still be a lack of affirmation due to false or inadequate results. Chest X-ray proves to be an excellent method, yet there is room for improvement in outcome accuracy. Computed tomography (CT) chest scans are processed through multiple stages to detect and narrow down the damaged region with the help of AI techniques [14, 15]. Some recent studies showed promising results in the early detection of COVID-19 signs. Narin et al. [12] have used three transfer learning models on the chest X-ray dataset to detect COVID-19 symptoms. They have achieved maximum accuracy in most of the folds, which can be an indication that the models may have shown overfit results. Wang et al. [16] fed CT images to a deep learning model, which can reveal damaged areas and extract required features that may help in diagnosing the disease. Shan et al. [17] used deep learning to develop a system that would automatically make multiple segments of lungs and reveal the infection.

Artificial intelligence and neural networks are being used readily in medicine to predict these kinds of viral diseases earlier through common symptoms [18]. Convolutional neural network (CNN), which is a type of deep learning, uses the images to train deep models and classifies them based on output categories [19]. This study is being used by researchers to train some state-of-the-art open-source neural network models and classify COVID-19 images. This branch of CNN is called transfer learning [20].

This paper suggests a deep learning approach to anticipate COVID-19 symptoms in a patient with the help of chest X-ray scans. In this study, transfer learning techniques were preferred over other machine learning algorithms due to the excellent classification accuracy of pretrained models, which also save time by avoiding the trouble of training and verifying the model weights from scratch. We have used five state-of-the-art predesigned networks in this study, including AlexNet, MobileNetv2, ShuffleNet, SqueezeNet, and Xception. These networks were fine-tuned by freezing most of the top convolutional layers and fully connected layers. Through multiple experiments on acquired datasets, we observed that the significant portion of the transfer learning model relies on the last convolutional layer for feature extraction and last fully connected layer for classification, which cannot be generalized for every dataset. Hence, only these layers were allowed to train weights. This approach not only saved a good amount of time but also provided competitive output accuracy. Moreover, each model is trained and tested on multiple optimizers as well as numerous learning rates to nullify the generalization factor, which is a crucial issue when there is a small amount of input data. The main highlights of the article are pointed as follows:

- (i) Five state-of-the-art transfer learning models are used with a fine-tuning approach to reduce training time while keeping the output accuracy intact
- (ii) Each model is trained multiple times at different learning rates to reassure that the models are not overfitting and showing false results
- (iii) MobileNetv2, when trained with the correct optimizer, provided the best results for chest X-ray images, although it is generally designed for mobile devices operations
- (iv) The proposed method is effective and robust due to the verification of several statistical measures obtained with a 10-fold cross-validation approach

The rest of our paper is as follows: Section 2 describes the datasets used in this study, Section 3 explains our proposed methodology, Section 4 examines the analytics metrics, Section 5 exhibits the results against both datasets, Section 6 discusses and compares our study with previous research while Section 7 concludes the paper.

2. Materials

For this study, two datasets were used to validate the transfer learning models' efficiency on X-ray images. The first dataset [21] contains 74 "normal" and 74 "pneumonia" images for training. It was taken from GitHub. 20 "normal" and 20 "pneumonia" images were used to test the integrity of models. The same number of images was used for the second dataset, where "normal" scans were taken from [21], while infected "pneumonia" ones were acquired from another open source [22]. The datasets were acquired from public source collection. Datasets are being updated on a regular basis, so the number of collected images may differ in future studies. It is shown in Table 1 that different pretrained models take different input sizes. Hence, all images were resized according to each pretrained model requirement training.

3. Methodology

This paper suggests an approach to detect COVID-19 in patients via chest X-ray scans. The proposed method contains three stages. The first stage works on preprocessing that was done on the data after it was obtained from open-source collection. Multiple images in the dataset had a different number of channels, so they could not be processed in model training. Initially, all images were converted to the same number of channel, that is, 3 in our case. As our input data are not big enough, to ensure that we get good output results, different data augmentation techniques, including rotating, image flip, and pixel change, have been used. During the second stage, different training parameters like the number of epochs, optimizer selection, number of folds per epoch, mini-batch size, learning rate, and model were defined. Data were resized distinctly according to each model. The third stage performs the classification step in which the network decides whether there are COVID-19 symptoms in the scans. As detection of this disease is a sensitive case,

TABLE 1: Key features of the models used in this study.

Model	Input size	Number of layers	Parameters (millions)	Size (MB)
AlexNet	$227 \times 227 \times 3$	8	61.0	227
MobileNetv2	$224 \times 224 \times 3$	53	3.5	13
ShuffleNet	$224 \times 224 \times 3$	50	1.4	6.3
SqueezeNet	$227 \times 227 \times 3$	18	1.24	4.6
Xception	$299 \times 299 \times 3$	71	22.9	85

numerous test runs were carried out to monitor the validity of trained models. Each transfer learning model was trained with three different optimizers, i.e., Adam, SGDM, RMSProp, and four learning rates, i.e., $1e-4$, $2e-4$, $3e-4$, and $4e-4$, to find the best combination and eliminate the factor of overfitting in data training. A flowchart of our suggested approach is shown in Figure 1.

3.1. Data Preprocessing. Both datasets were processed through two stages to endorse the maximum output accuracy. Due to the presence of different numbers of channels for different images, all images were converted to the same channel size during the first stage. Neural network models require substantial data for training. As our input data was not large enough, data augmentation was performed to ensure that each model fed on enough input images to avoid overfitting. Data augmentation is a process in which images are modified by applying small changes in the original pictures like rotation, flipping the image, and minor adjustment in pixel range. An example of data augmentation is shown in Figure 2.

3.2. Transfer Learning. Deep learning focuses on functioning as a human mind. When a child is taught about different animals, an arbitrary image is formed in the mind of the child that a dog looks like this and a cat looks like this, and in the future, the child can recognize these animals. Deep learning works on the same principle. Transfer learning is the next step in deep learning. Training a neural network model requires a lot of time and multiple runs to capture the accurate weights according to the model's requirement. It is a tedious work and is not easy for students new to the field to enter transfer learning. Transfer learning handles the models shared by field experts for the public, which skips the requirement of finding compatible weights and carries on to the next step of the training model on new input data. We have used following pretrained models in our study:

- (i) AlexNet
- (ii) MobileNetv2
- (iii) ShuffleNet
- (iv) SqueezeNet
- (v) Xception

Figure 3 gives visual on the architecture of these models. The blue block defines the input. Yellow indicates the convolutional layer. Orange box performs rectified linear unit (ReLU) operation. Green is responsible for cross

channel normalization. The purple box is used for normalization, gray box concatenates the above results, and the white box represents channel shuffling. Some functions are unique to some models. For instance, MobileNetv2 executes clipped ReLU operation instead of general ReLU, and ShuffleNet contains multiple grouped convolutional layers.

3.2.1. AlexNet. Given less number of computational parameters, as compared to other models with comparable performance for nearly every data form, AlexNet is one of the most famous pretrained models among researchers. It has five convolutional layers and three fully connected layers for classification purposes [23]. We only used the last convolution layer for feature extraction and the final fully connected layer for classification to reduce training time while keeping the accuracy unscathed. The image input size for AlexNet is $227 \times 227 \times 3$.

3.2.2. MobileNetv2. Originally designed for mobile devices by Google, MobileNetv2 is a fine pretrained model that delivers high output accuracy. It is designed to work with low input resources and reduced mathematical calculations. The working principle of this model is depthwise separable convolution and linear bottlenecks [24]. The second version of MobileNet, or as we call it MobileNetv2, also introduced short connections between bottlenecks. The input size of this model is $224 \times 224 \times 3$.

3.2.3. ShuffleNet. ShuffleNet is an extremely efficient model of a convolutional neural network that was also initially designed for mobile devices. It has an impressive computational power of 10–150 MFLOPs. This model operates on pointwise group convolution. ShuffleNet works on channel shuffling to reduce computational parameters and achieve high output accuracy. The model has proven work better than the “MobileNet” system for the classification of images [25]. It takes an input frame of the size $224 \times 224 \times 3$.

3.2.4. SqueezeNet. This 18-layer deep convolutional network was designed to achieve similar accuracy as AlexNet with 50x fewer parameters to compute. The idea behind reducing computation parameters is through the replacement of 3×3 filters with 1×1 filters. By performing this small operation, the model requires nine times calculations to perform. Another important concept of SqueezeNet is “Fire Module”. The squeezed layer feeds into the expand layer (3×3 filter) to reduce filter size and hence calculations. This architecture is

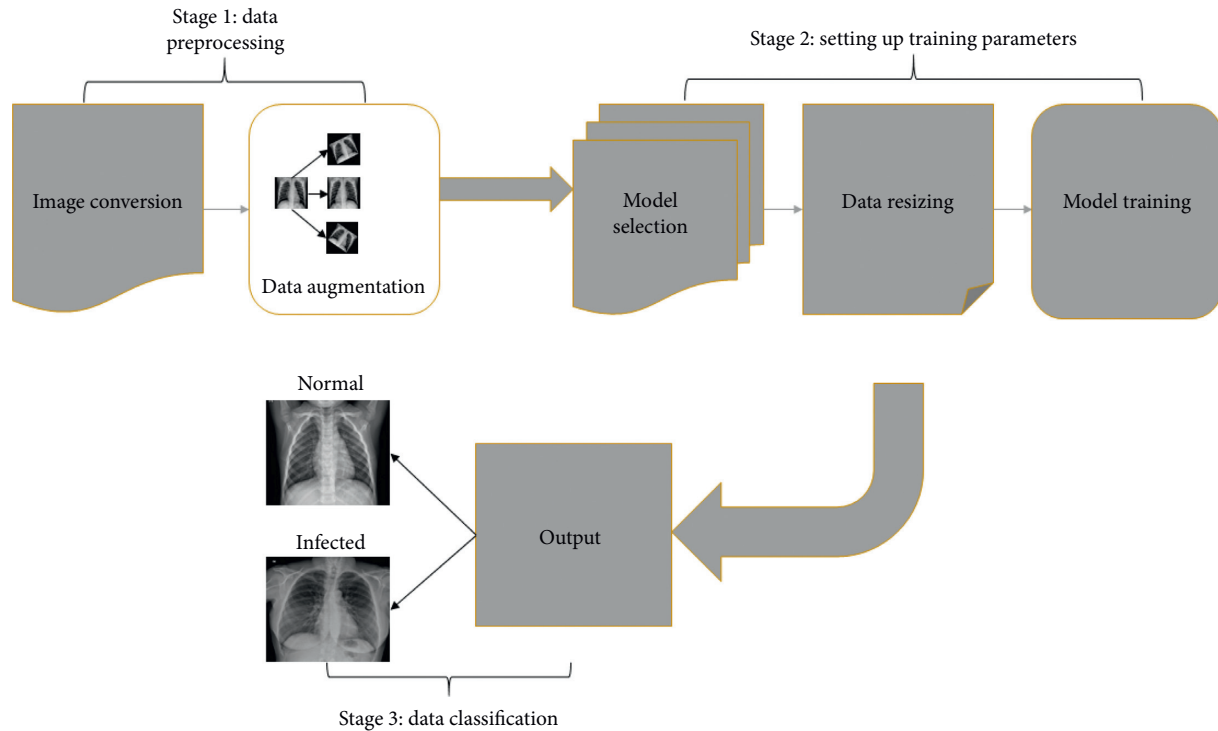


FIGURE 1: Block diagram of our study. Stage 1: data preprocessing: the number of channels of different images was made alike and data augmentation was performed. Stage 2: setting up training parameters: training parameters like number of epochs, mini-batch size, and number of folds per epoch opted in this stage. Data training was performed on each model after image resizing according to the distinct input size. Stage 3: data classification: here, our trained model displays the classified result as either normal or infected.

known as “Fire Module” [26]. SqueezeNet has a similar input size as AlexNet.

3.2.5. Xception. Xception or extreme version of the inception model is a pretrained neural network that operates on modified depthwise separable convolution. In simple depthwise convolution operation, channelwise $n \times n$ spatial convolution is being performed, while in the modified version (Xception), pointwise convolution is followed by depthwise convolution [27]. Xception has outperformed VGG, ResNet, and Inception-v3 in ImageNet competition. The input size of Xception is $299 \times 299 \times 3$. Some key features of these models are presented in Table 1.

3.3. Matlab Application. An application related to our study was designed in the graphical user interface environment (Matlab GUI) of Matlab 2019b. It will assist researchers in the future study for coronavirus detection through chest X-ray images. Matlab app is a built-in program that is used to automate the required task. Multiple test runs were performed to corroborate the image ranking and run-time of the App. To get the best possible classification outcome, the finest transfer learning models used in this analysis, including “MobileNetv2” (against two different optimizers), “SqueezeNet”, and ‘Xception’, were incorporated into the app. These networks have provided excellent classification results on chest X-ray images in our study. Figure 4 portrays the app’s function. It will take single image input and

labeling the picture into one of our study-focused categories, i.e., normal or infected. Following are the components and their functions embedded in the app design:

- (1) Input: load the input data from the folder. Input will be in the form of an image; hence, the app is designed to accept file format of “.png”, “.jpg,” and “.jpeg”.
- (2) Model Selection: requires the user to select either of four models integrated into the framework of the app, where Model 1 and Model 2 represent MobileNetv2 (RMSProp, LR: $3e-4$) and SqueezeNet (Adam, LR: $3e-4$) trained with the dataset with 1 image while Model 3 and Model 4 demonstrate MobileNetv2 (Adam, LR: $3e-4$) and Xception (RMSProp, LR: $3e-4$) trained with the dataset with 2 images.
- (3) Axes: display the classified image, i.e., if the models detect no symptoms of COVID-19, the image will be displayed in the “normal” axis; otherwise, it will be portrayed in “infected” axis.

Parameters like epochs, optimizer, learning rate, and mini-batch size are preselected in the models as it is required to use integrated models in the app.

4. Performance Measures

We have used a built-in Matlab deep learning toolbox to train different transfer learning models on our input data. Each model was trained using the 10-fold procedure to

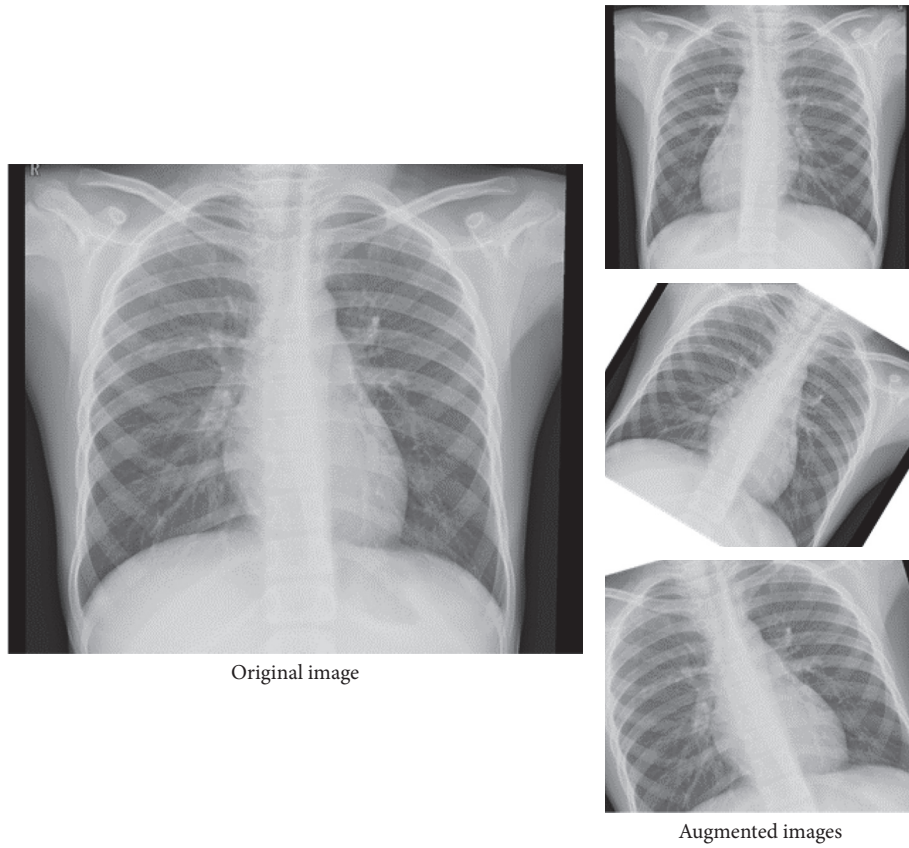


FIGURE 2: A visual representation of data augmentation.

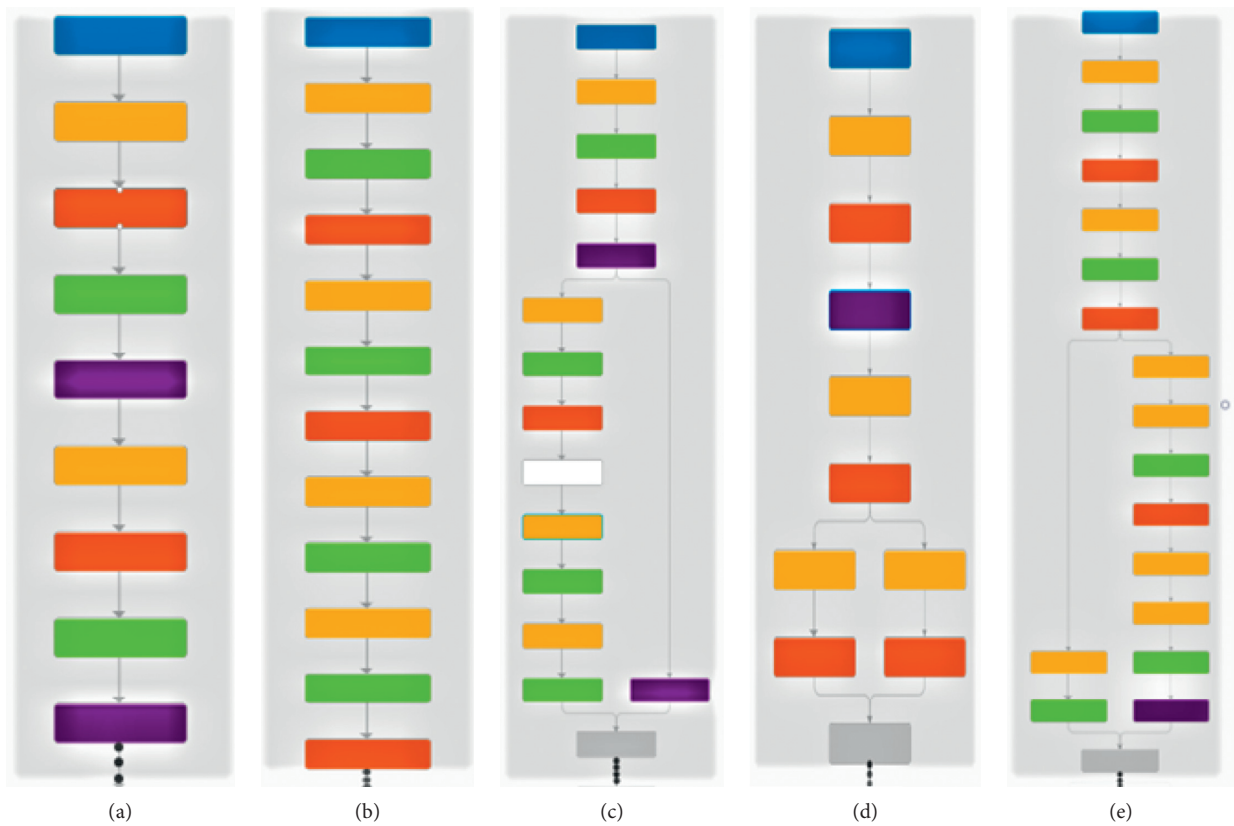


FIGURE 3: Frameworks of pretrained models: (a) AlexNet, (b) MobileNetV2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

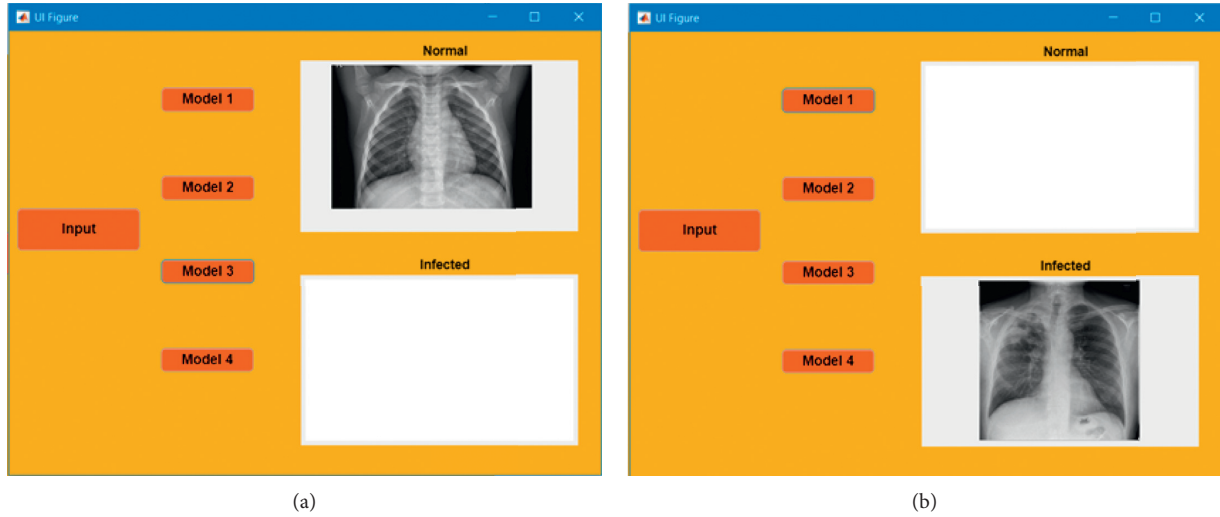


FIGURE 4: The interface of Matlab App is designed according to this study. (a) indicates when the model(s) classify the input image as to normal, while (b) symbolizes infected or pneumonia class.

guarantee the validity of the result. Each training run consisted of 10 epochs and 140 iterations per epoch. Models were fine-tuned before training. All the layers except the last one were frozen to avoid extra time consumption. The classification layer and the final fully connected layer of each model were replaced as they were originally designed to provide an output of 1000 distinctive categories. In each fold, out of 74 images, 15 images were separated randomly for validation tests.

All the training and test simulations were performed on an Intel Core i7-9750H processor enforced with 32 gigabytes of RAM and the GPU (graphics processing unit) Nvidia GeForce GTX 1660 Ti with 6 GB memory. Matlab framework was restarted before each new training to assure there is no false time consumption, which can incur when an excessive number of intense simulations are executed.

Detection of coronavirus among healthy people has become one of the top priorities of doctors worldwide. Results generated through these methods must be validated via multiple techniques because any false result can be very dangerous not only to that patient but also to other people in contact with that patient. All models were validated through analytics metrics, including overall accuracy, precision, recall, and F -score. The following equations represent the mathematical formulas of these metrics:

$$\begin{aligned}
 \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN}, \\
 \text{Recall} &= \frac{TP}{TP + FN}, \\
 \text{Precision} &= \frac{TP}{TP + FP}, \\
 F - \text{measure} &= \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}},
 \end{aligned} \tag{1}$$

where TP is “true positive”, TN is “true negative”, FP is “false positive,” and FN is “false negative,” respectively. These parameters are used to analyze the integrity of test results [28, 29]. Accuracy is the measurement of correctly classified samples in percentage or closeness of the measured value to a standard or true value. The number of positive class predictions from all positive examples in the dataset is defined as recall. Precision is the ratio of positive observations correctly predicted to the overall positive observations predicted, while F -measure gives a mean for both precision and recall to be integrated into a single measure that captures both properties. It is the harmonic mean of precision and recall.

5. Results

This study was carried out to diagnose patients with COVID-19 symptoms with the help of chest X-ray scans. Various deep learning models were trained and tested on multiple optimizers and several learning rates. The reason for performing this study on numerous parameters is to find the optimum combination of model, optimizer, and learning rate for our input data.

5.1. Dataset 1. Accuracy comparison of the first dataset is shown in Figure 5. We can see that MobileNet2 adopted all three optimizers very well for all learning rates except for 1e-4, which is not uncommon in other models. MobileNet2 synthesized the highest accuracy of 97% with “Adam” optimizer at a learning rate of 3e-4. ShuffleNet showed mixed results with a maximum output efficiency of 89% on two different combinations. For SGDM optimizer, Xception has shown surprisingly bad results falling up to 60% of average accuracy with LR=1e-4, which is worse than all other scenarios in our study. SqueezeNet showed prominent results for all learning rates against different optimizers,

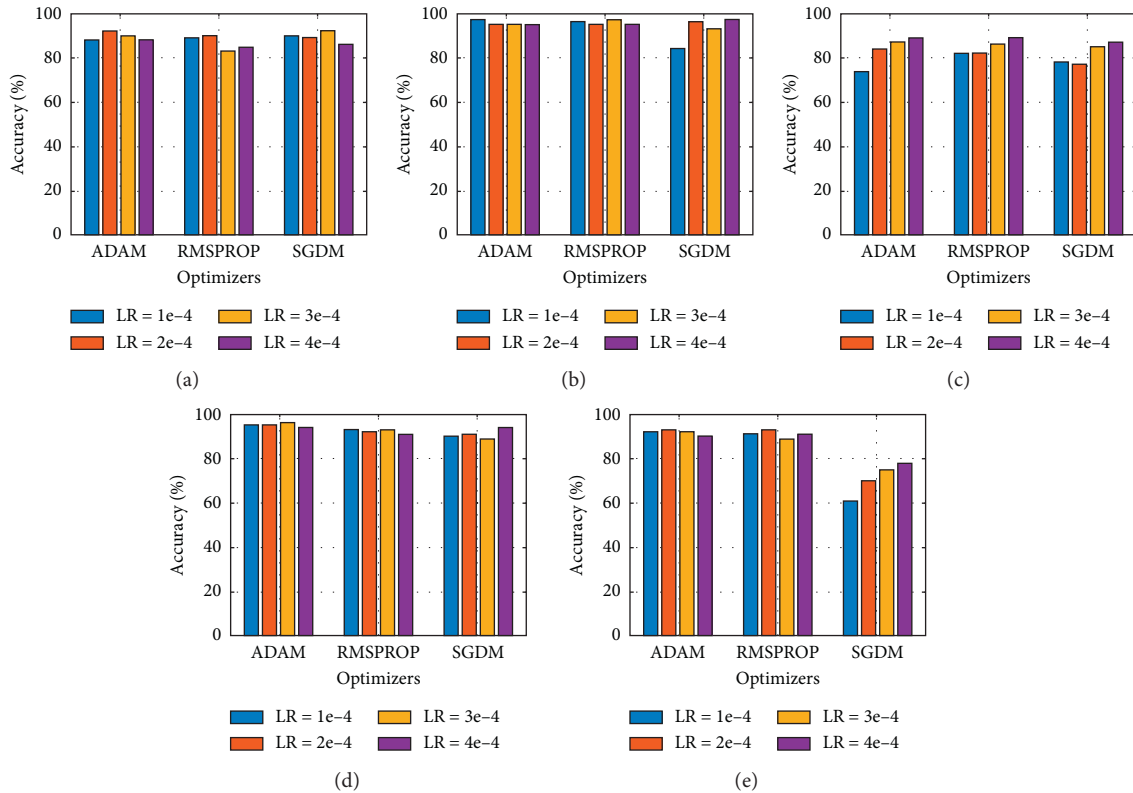


FIGURE 5: Graphical comparison of different models based on output accuracy for dataset 1: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

reaching a maximum of 96% classification accuracy for Adam optimizer when the learning rate was selected as 3e-4.

Figure 6 depicts the training time of these models in different cases. On average, AlexNet has taken the least amount of time for each training except when trained with RMSPROP. MobileNetv2 expressed diverse training time for different optimizers. ShuffleNet and SqueezeNet both registered the maximum amount of computational time for all three optimizers, nearly approaching 100 seconds per run. AlexNet only consists of 8 layers, which is far less as compared to that of the other four models. So, its less time consumption is understandable, but MobileNetv2 results were somewhat surprising, taking far less training time and showing excellent classification results. Xception, as expected, required maximum training time as it is one of the most in-depth networks used in our study.

All in all, almost every model has adopted well with SGDM according to time usage. If we compare Figure 4 and Figure 5, we can quickly notice that SqueezeNet with Adam optimizer is probably the best combination of both accuracy and time consumption.

Confusion matrices for dataset 1 are shown in Figure 7, while Table 2 exhibits a comparison of precision, recall, and F -score for dataset: 1, where MobileNetv2 has attained the best F -score for both “normal” and “infected” classes. Again, “infected” here represent patients who showed pneumonia signs during medical tests. MobileNetv2 also got the highest precision score of 98%

for “infected”, which is on par with Xception for the same case.

5.2. Dataset 2. As mentioned before, we have used two datasets to verify the integrity of models. The following data re half part of the first dataset and the other half is extracted from another source [22]. Exact operations were performed on dataset:2 as were on dataset 1. A comparison of average accuracy for different models is given in Figure 8. Xception has shown a similar pattern here with SGDM. So, it is not recommended to use Xception for this dataset classification with either optimizer. Results can be improved with a big dataset as the Xception model works best on substantial data size, that is, if you want to use SGDM optimizer with Xception. However, a maximum result of 96% with RMSProp at LR: 3e-4 is still acquired, the best classification accuracy for dataset 2. An inclined configuration can be seen for MobileNetv2 when used with SGDM where output accuracy showed a direct relation with the learning rate, and it peaked at 95%.

However, with the other two optimizers, the results are very good for MobileNetv2. SqueezeNet has also produced excellent results with Adam as well as with RMSProp marking up to 94% output accuracy. The time consumption graph for each model against dataset 2 is shown in Figure 9.

Though Xception took the least amount of training time with this dataset, it is not recommended to use due to

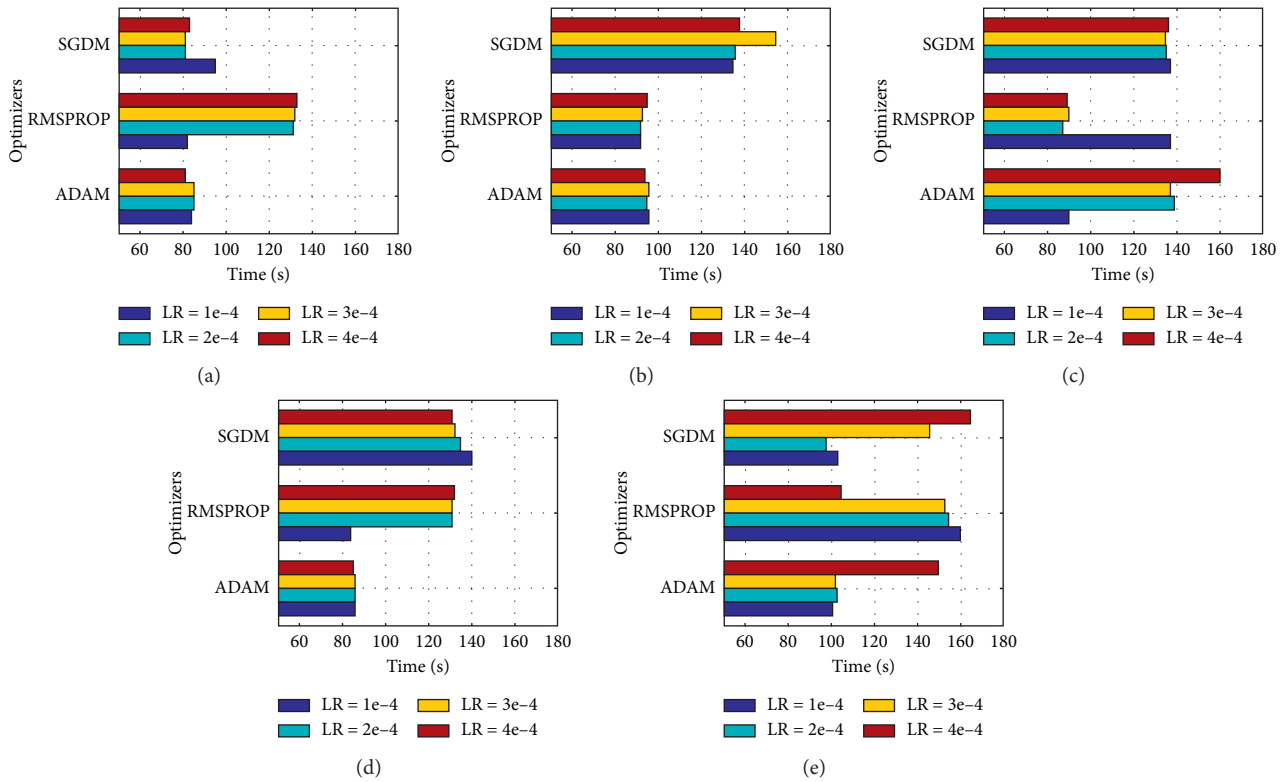


FIGURE 6: Time comparisons of all models for dataset 1: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

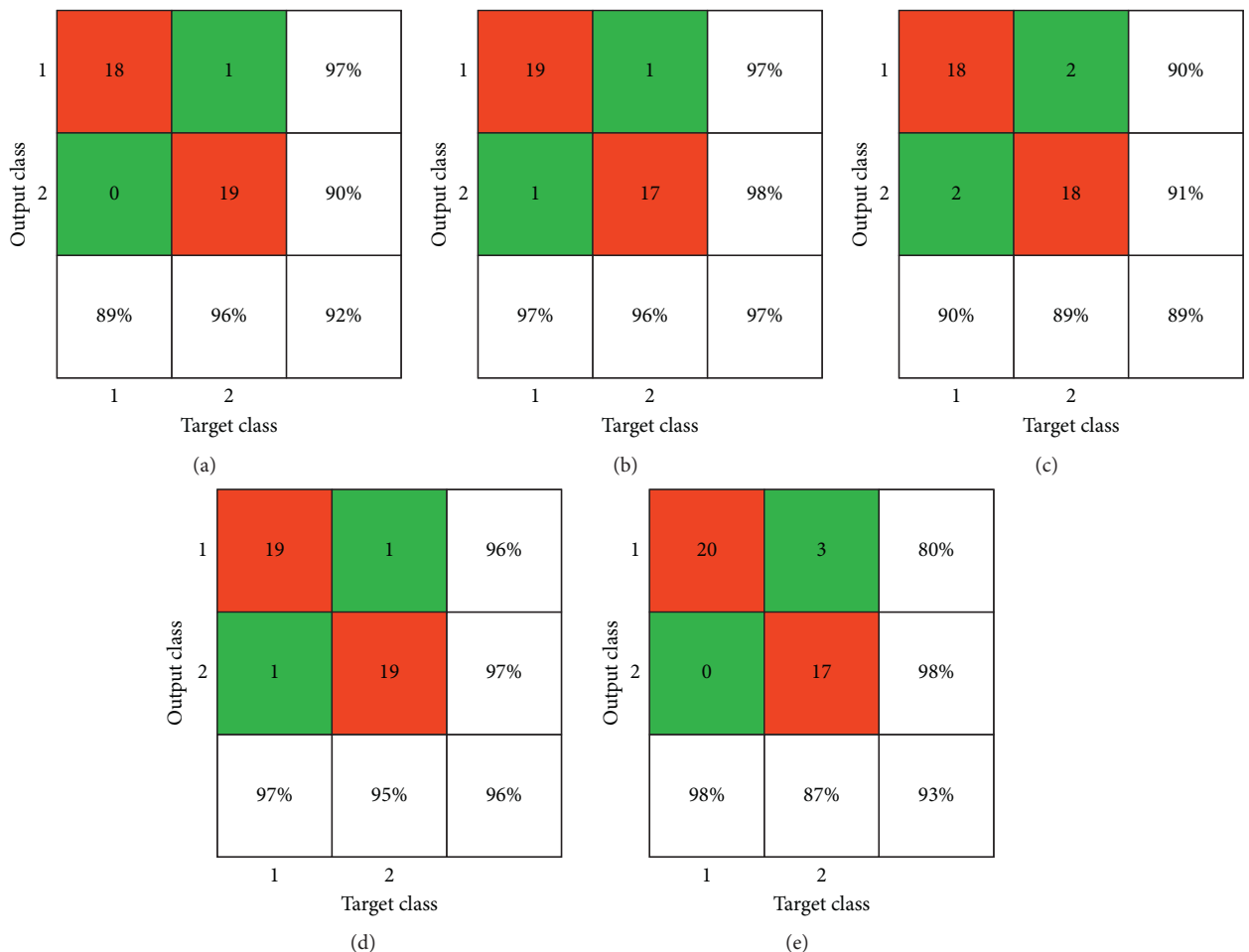


FIGURE 7: Confusion matrices of experimented models for dataset 1: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

TABLE 2: Statistical measurement comparison of observed models for dataset 1.

Network	Class	Recall	Precision	F-score
AlexNet	Normal	89	96	92
	Infected	96	90	93
MobileNetv2	Normal	97	97	97
	Infected	96	98	97
ShuffleNet	Normal	90	90	90
	Infected	89	91	90
SqueezeNet	Normal	97	96	96
	Infected	95	97	96
Xception	Normal	98	89	93
	Infected	87	98	92

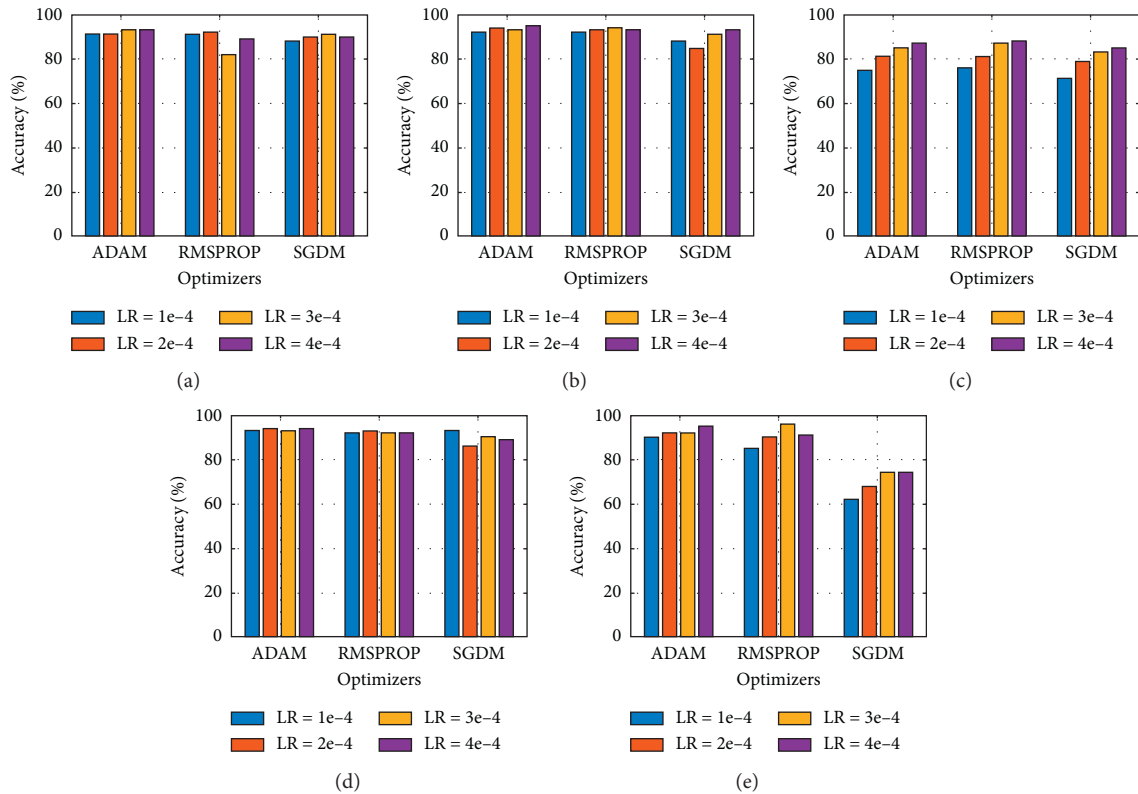


FIGURE 8: Graphical comparison of different models based on the output accuracy for dataset 2: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

significantly less output accuracy. Other models show more or less similar results with little fluctuation where time is taken into account.

Figure 10 represents the confusion matrices of all models used in this study for the second dataset. Remarkably, Xception synthesized 96% accuracy as the best one; still, it did not show good average result. MobileNetv2 was second-best, which yielded 95% output accuracy.

Table 3 indicates that Xception attained the finest result with “normal” class while calculating recall, and it also measures 100% precision for “infected” class. That is why the F-score of Xception was the best among all the models. Table 3 and Figure 8 represent the best result that we observed. If we talk about average output, MobileNetv2 seems to be the clear winner.

6. Discussions

Numerous studies have been performed on the detection of COVID-19 symptoms via different techniques. Shan et al. [17] used VB-net for the image segmentations of patients. A study similar to ours was conducted in [12] where they achieved 98% accuracy. But, their results could be prone to overfitting as they did not use multiple optimizers or different learning rates and only used three transfer learning methods. Zhang et al. [30] performed X-ray images classification with the help of ResNet. Wang and Wong [31] adopted a convolutional neural network method for the classification of X-ray images. They successfully achieved 83.5% accuracy. A very famous transfer learning model “inception” was used by Wang et al. [16] to predict COVID-

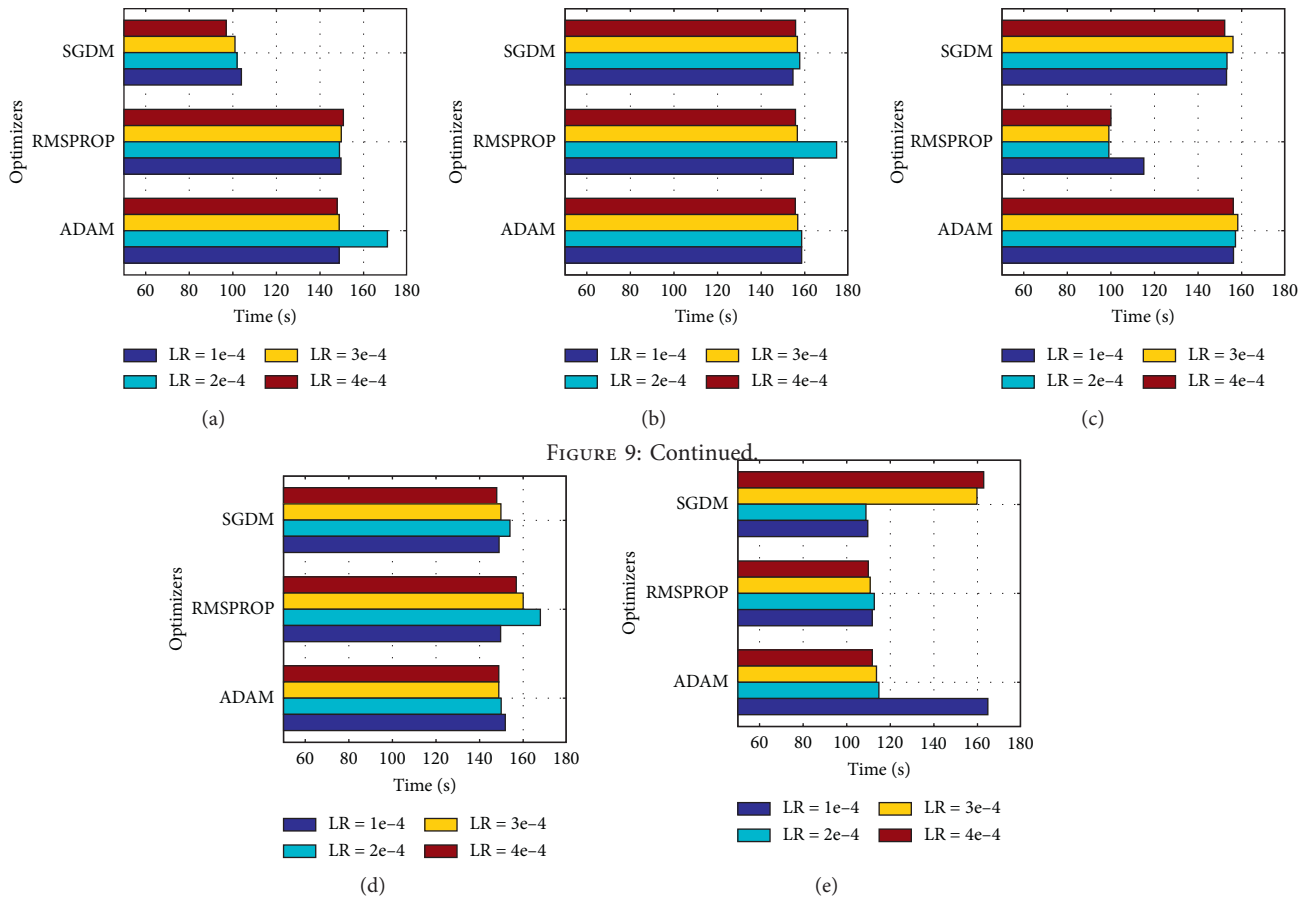
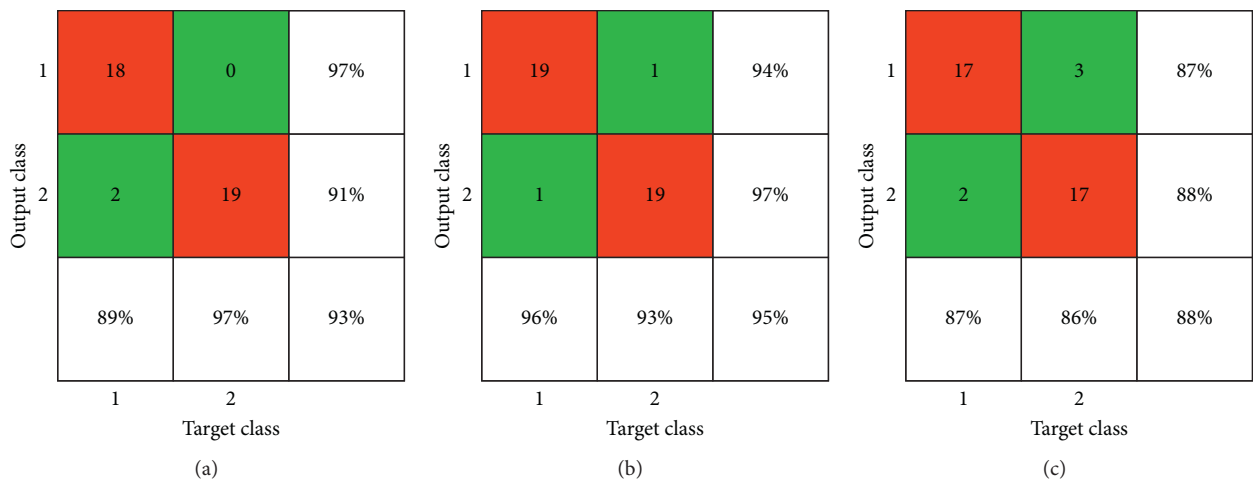


FIGURE 9: Time comparisons of all models for dataset 2: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.



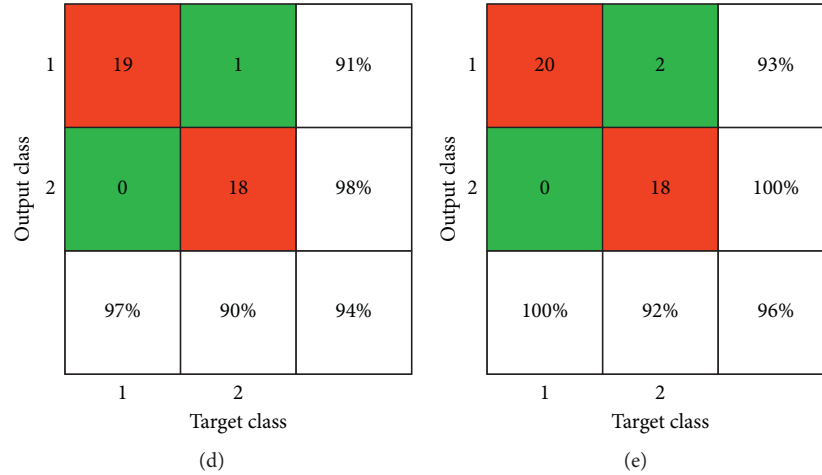


FIGURE 10: Confusion matrices of experimented models for dataset 2: (a) AlexNet, (b) MobileNetv2, (c) ShuffleNet, (d) SqueezeNet, and (e) Xception.

TABLE 3: Statistical measurement comparison of observed models for dataset 2.

Network	Class	Recall	Precision	<i>F</i> -score
AlexNet	Normal	89	97	93
	Infected	97	91	94
MobileNetv2	Normal	96	94	95
	Infected	93	97	95
ShuffleNet	Normal	87	87	93
	Infected	86	88	92
SqueezeNet	Normal	97	91	94
	Infected	90	98	94
Xception	Normal	100	93	96
	Infected	92	100	96

19 symptoms in CT images. The majority of these studies have performed classification on fewer neural network models as compared to our research. Furthermore, we have conducted training on different optimizers as well as on different learning to confirm that there is no overfitting going on due to the lack of big datasets. Our method is rigorous and repetitive, as we have performed 10-fold cross-validation. Table 4 provides a quick overview of our findings as compared to several other studies that used similar datasets for neural network model training. Also, after carrying out numerous additional simulations, we have achieved near-best precision to ensure that the results produced are not false or due to a computational error. Our analysis is highlighted in the following points:

(i) Multiple transfer learning models, including AlexNet, MobileNetv2, ShuffleNet, SqueezeNet, and Xception, have been used to classify chest X-ray images with different optimizers and learning rates to synthesize accurate results

(ii) Fine-tuning has been used to reduce the computational parameters and make use of only those layers which take part in feature extraction

TABLE 4: Comparison of different studies.

Paper	Dataset	Objective	Approach	Highest avg. accuracy
Ghoshal et al. [32]	X-ray	COVID-19-image classification	CNN	92.9%
Pan et al. [8]	X-ray	COVID-19-image classification	ResNet50	98.0%
Zhang et al. [30]	X-ray	COVID-19-image classification	ResNet	95.18% (AUC)
Wang et al. [16]	X-ray	COVID-19-image classification	CNN	83.5%
Our paper	X-ray	COVID-19-image classification	MobileNetv2	97.0%

(iii) We have used X-ray images, which are not difficult to acquire, and showed that they could be beneficial in the detection of COVID-19 in a patient

(iv)The problem with limited input data has been solved by making use of different data augmentation techniques

This study was just one way to diagnose COVID-19 symptoms in patients. Several other transfer learning models can be beneficial in image classification. Moreover, deep learning always relies on the amount of input data. Hence, if a large amount of data can be collected, it will further assist in getting enhanced results. For instance, Xception with SGDM showed relatively poor results due to insufficient input data.

7. Conclusion

Because of its fast-spreading potential, COVID-19 has rapidly become the key target of doctors and medical researchers around the world. It is critical to detect this virus in humans in the absence of a functional vaccine to prevent its dissemination. This paper emphasizes on using chest X-ray scans to diagnose COVID-19 symptoms. The proposed study implements five different transfer learning models with different optimizers and various learning rates on two public datasets. Results dictate that MobileNetv2 and Xception models can be instrumental in diagnosing coronavirus through chest X-ray images. To authenticate the effectiveness and robustness of trained model, all models were validated by several statistical indexes, including a 10-fold cross-validation method. We believe that this study can be a big help in the early detection of COVID-19.

Data Availability

The data used to support the findings of this study are available upon request by contacting the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] K. Roosa, Y. Lee, R. Luo et al., "Real-time forecasts of the COVID-19 epidemic in China from february 5th to february 24th, 2020," *Infectious Disease Modelling*, vol. 5, pp. 256–263, 2020.
- [2] L. Yan, H.-T. Zhang, J. Goncalves et al., "A machine learning-based model for survival prediction in patients with severe COVID-19 infection," *medRxiv*, pp. 1–18, 2020.
- [3] S. B. Stoecklin, P. Rolland, Y. Silue et al., "First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, january 2020," *Eurosurveillance*, vol. 25, no. 6, Article ID 2000094, 2020.
- [4] WHO, "Coronavirus disease 2019 (COVID-19) situation report-76," 2020, <https://www.who.int/docs/default-source/coronavirus/situation-reports/20200405-sitrep-76-covid-19.pdf>.
- [5] M. T. Sadiq, X. Yu, and Z. Yuan, "Exploiting dimensionality reduction and neural network techniques for the development of expert brain-computer interfaces," *Expert Systems with Applications*, vol. 164, Article ID 114031, 2020.
- [6] <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>.
- [7] C. Zheng, X. Deng, Q. Fu et al., "Deep learning-based detection for COVID-19 from chest CT using weak label," *medRxiv*, 2020.
- [8] F. Pan, T. Ye, P. Sun et al., "Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia," *Radiology*, vol. 295, pp. 715–721, Article ID 200370, 2020.
- [9] J. P. Kanne, "Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist," *Radiology*, vol. 295, pp. 16–17, Article ID 200241, 2020.
- [10] A. Bernheim, X. Mei, M. Huang et al., "Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection," *Radiology*, vol. 295, Article ID 200463, 2020.
- [11] M. T. Sadiq, X. Yu, Z. Yuan, and M. Z. Aziz, "Identification of motor and mental imagery EEG in two and multiclass subject-dependent tasks using successive decomposition index," *Sensors*, vol. 20, no. 18, p. 5283, 2020.
- [12] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-Ray images and deep convolutional neural networks," 2020, <https://arxiv.org/abs/2003.10849>, Article ID 10849.
- [13] <https://www.nationalgeographic.com/science/2020/02/here-is-what-coronavirus-does-to-the-body/20.03.2020>.
- [14] T. Ai, Z. Yang, H. Hou et al., "Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, Article ID 200642, 2020.
- [15] M. T. Sadiq, X. Yu, Z. Yuan et al., "Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain-computer interfaces," *IEEE Access*, vol. 7, pp. 171431–171451, 2019.
- [16] S. Wang, B. Kang, J. Ma et al., "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *medRxiv*, pp. 1–26, 2020.
- [17] F. Shan, Y. Gao, J. Wang et al., "Lung infection quantification of COVID-19 in CT Images with deep learning," pp. 1–19, 2020, <https://arxiv.org/abs/2003.04655>.
- [18] H. Nazeran, F. Rice, W. Moran, and J. Skinner, "Biomedical image processing in pathology: a review," *Australasian Physical & Engineering Sciences in Medicine*, vol. 18, no. 1, pp. 26–38, 1995.
- [19] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 253–256, Paris, France, May 2010.
- [20] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [21] Kaggle, "Chest X-Ray images (pneumonia)," <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- [22] Open Database of COVID-19 Cases with chest X-Ray or CT images, <https://github.com/ieee8023/covid-chestxray-dataset>.
- [23] M. Z. Alom, T. M. Taha, C. Yakopcic et al., "The History began from AlexNet: a comprehensive survey on deep learning approaches," <http://arxiv.org/abs/1803.01164>.
- [24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: inverted residuals and linear bottlenecks," 2018, <https://arxiv.org/abs/1801.04381>.
- [25] X. Zhang, X. Zhou, M. Lin, and J. S. Shufflenet, "An extremely efficient convolutional neural network for mobile devices," 2017, <https://arxiv.org/abs/1707.01083>.

- [26] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: alexnet-level accuracy with 50x fewer parameters and <<1mb model size," 2016, <https://arxiv.org/abs/1602.07360>.
- [27] F. Chollet, "Xception: deep learning with depthwise separable convolutions," 2016, <https://arxiv.org/abs/1610.02357>.
- [28] M. Junker, R. Hoch, and A. Dengel, "On the evaluation of document analysis components by Recall, precision, and accuracy," in *Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR '99)*, pp. 713–716, Bangalore, India, 1999.
- [29] D. M. W. Powers, "Evaluation: from precision, recall and F-factor to ROC, informedness, Markedness & Correlation," Technical report SIE-07-001, Flinders University, Adelaide, Australia, 2007.
- [30] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 screening on chest X-ray images using deep learning based anomaly detection," vol. 12338, 2020.
- [31] L. Wang and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from Vhest radiography images," 2020, <https://arxiv.org/abs/2003.09871>.
- [32] B. Ghoshal and A. Tucker, "Estimating Uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection," p. 2020, 2003, <https://arxiv.org/abs/2003.10769>, Article ID 10769.

Research Article

Surgical Design Optimization of Proximal Junctional Kyphosis

Li Peng,¹ Guangming Zhang,¹ Heng Zuo,^{1,2} Lan Lan ,¹ and Xiaobo Zhou³

¹West China Biomedical Big Data Center, West China Hospital/West China School of Medicine, Sichuan University, Chengdu 610041, China

²School of Mathematics, Sichuan Normal University, Chengdu 610066, China

³Center for Computational Systems Medicine, School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston 77030, USA

Correspondence should be addressed to Lan Lan; lanl@scu.edu.cn

Received 7 July 2020; Revised 19 August 2020; Accepted 1 September 2020; Published 21 September 2020

Academic Editor: Antonio Gloria

Copyright © 2020 Li Peng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Purpose. The objective of this study was to construct a procedural planning tool to optimize the proximal junction angle (PJA) to prevent postoperative proximal junctional kyphosis (PJK) for each scoliosis patient. **Methods.** Twelve patients (9 patients without PJK and 3 patients with PJK) who have been followed up for at least 2 years after surgery were included. After calculating the loading force on the cephalad intervertebral disc of upper instrumented vertebra of each patient, the finite-element method (FEM) was performed to calculate the stress of each element. The stress information was summarized into the difference value before and after operation in different regions of interest. A two-layer fully connected neural network method was applied to model the relationship between the stress information and the risk of PJK. Leave-one-out cross-validation and sensitivity analysis were implemented to assess the accuracy and stability of the trained model. The optimal PJA was predicted based on the learned model by optimization algorithm. **Results.** The mean prediction accuracy was 83.3% for all these cases, and the area under the curve (AUC) of prediction was 0.889. And the output variance of this model was less than 5% when the important factor values were perturbed in a range of 5%. **Conclusion.** Our approach integrated biomechanics and machine learning to support the surgical decision. For a new individual, the risk of PJK and optimal PJA can be simultaneously predicted based on the learned model.

1. Introduction

For adolescent idiopathic scoliosis (AIS) patients, orthopedic operations are employed to reconstruct the coronal and sagittal alignment to maintain the stability of spine [1]. Long posterior instrumentation and fusion surgery is often a powerful surgical treatment for spinal deformity [2, 3]. During the treatment, vertebrae are fused using pedicle screws or other combinations of devices. Such fusion treatment is intended to reconstruct spinal geometry by strong correction and derotation of the spine [4]. However, the lack of mobility in fusion segments has raised a postulation that such fusion may increase the stress in proximally cephalad spinal segments and eventually accelerate deterioration of the neighboring discs [5].

Powerful correction maneuvers of predominantly all-pedicle instrumentation could also result in a series of issues

such as the increase of the proximal junction angle (PJA, the sagittal Cobb angle between the inferior endplate of the upper instrumented vertebra (UIV) and the superior endplate of two cephalad vertebrae) [6]. Proximal junctional kyphosis (PJK), an abnormal kyphotic deformity involving spinal segments proximally adjacent to the fusion segments, has drawn the attention of many spine surgeons [7–9]. This frequent complication might cause regional pain, diminish quality of life, and ultimately lead to revision surgery in some severe cases [10, 11]. The generally accepted definition of PJK is described by Glattes et al. [8] that PJA is more than 10° and at least 10° greater than the preoperative measurement. The incidence rate of PJK ranges extensively from 6.0% to 45.1% [6, 8, 10, 12–14]. A retrospective review of 836 adult cases reported a higher percentage of unplanned readmission due to PJK within 90 days from surgery (51.9%) compared with other surgical complications [15].

The recent advances in computer-aided design (CAD) have been rapidly changing the landscape in scoliosis treatment procedures, improving the clinical outcomes of patients significantly as valuable models are practiced [5, 16]. Many studies biomechanically assess and evaluate the independent effects of different instrumentation variables by the finite-element method (FEM) [15, 17–21]. The difficulty mainly comes from a finite range of optimizations by manual selection. Our study focuses on designing a reliable automatic system to assist surgeons designing and preoperatively help decrease the readmission rate of scoliosis patients.

The larger difference value of preoperative and postoperative PJA (more than 5°) has been determined as a risk factor of PJK [8]. However, the individual and coupling biomechanical effects of PJA are not yet fully understood. The purpose of this paper is to predict which AIS patients have higher risks of PJK due to biomechanical factors and calculate the optimal PJA for each AIS patient. In this work, we hypothesize that inappropriate intraoperative PJA change compared with preoperation may lead to inhomogeneous distribution of loading and ultimately result in various degrees of degeneration in the cephalad intervertebral disc of UIV, and a suitable angle is beneficial for an individual patient to have a better prognosis after corrective surgery. The biomechanical information of intervertebral discs in the proximal junctional segment following spinal deformity surgery can be accurately simulated by integrating a finite-element method (FEM) with a statistical learning model.

2. Materials and Methods

This study presented an integrated approach to accurately simulate cephalad intervertebral disc behavior of the UIV for pre- and postoperation, respectively, for the purpose of optimizing the PJA for AIS patient. Figure 1 describes the flowchart of the whole process.

This retrospective study was completed with AIS patients who undergone scoliosis correction surgery from 2013 to 2018 at West China Hospital and have been followed-up for at least 2 years to assess whether or not developing PJK. Patients were excluded if imaging information including preoperative, immediate postoperative (3–7 days after surgery) X-ray, and preoperative CT could not be obtained. Twelve cases were recruited for this biomechanical study. For each patient, the collected data involved the upper instrumented vertebral documentation, actual PJA, gender, age, and preoperative spinal computed tomography (CT). Based on the postoperative 2-year PJA, which is more than 10° and at least 10° greater than the preoperative measurement, those patients were categorized into 2 groups: PJK group and non-PJK group. In this study, we used the aforementioned study information for each patient as ground truth to confirm the patient selection.

2.1. Feature Extraction

2.1.1. Intervertebral Disc Segmentation and Quantification. Twelve subject-specific geometries of patients were acquired using preoperative CT scans. All images were acquired with

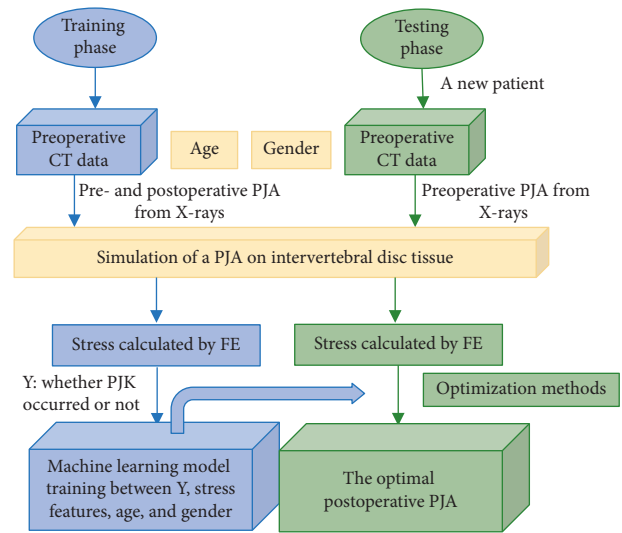


FIGURE 1: The flowchart of optimizing the surgical design of PJK. In the training phase, the machine learning model is generated. During the testing phase, the model can be used for PJA optimization based on a new patient's information.

1 mm slice intervals and a 512×512 acquisition matrix, and then imported into Mimics 20.0 Imaging Software (Materialise, Leuven, Belgium) for segmenting as shown in Figure 2. Initial segmentation was performed by thresholding image from 50 to 150 Hounsfield units chosen to most accurately preserve intervertebral disc geometry. Manual segmentation was employed to delineate regions, which were visible but could not be captured by automated methods. To limit the area-of-interest and reduce computational complexity, we restricted the zone to the cephalad intervertebral disc of UIV because no tissue deformations appeared in adjacent vertebrae (as shown in Figure 2(b)).

For spine corrective operation, a preprocedural plan is meaningful only if it can be accurately transferred to a patient at the time of intervention. For this reason, we applied a validated software (Surgimap, version 2.2.15.5), which could measure the degree of the curvature quickly on pre- and postprocedural standing radiographs, thus determining the influence of gravity at each vertebral level of patients with scoliosis in the upright position [22]. The severity of scoliosis can be evaluated by measuring the Cobb angle. With using strong correction of all-pedicle instrumentation, the PJA changes with reconstructing the coronal and sagittal alignment [6]. After selecting the most suitable UIV according to the scoliosis type, assuming different stress distribution of the UIV surface would derive from angular variations of PJA, and inappropriate intraoperative PJA change compared with preoperation could lead to inhomogeneous stress distribution of the upper body weight on the UIV surface and the degeneration of adjacent intervertebral disc. Usually, a virtual proximal segment correction will decrease the PJA as possible. However, due to the biomechanical properties of tissues, patients may still be at risk for PJK two years after spinal surgery because of the nonspecific PJA. Therefore, our approach addresses the need to develop a reliable process for simulating tissue behavior changes of intervertebral disc between pre- and postprocedure.

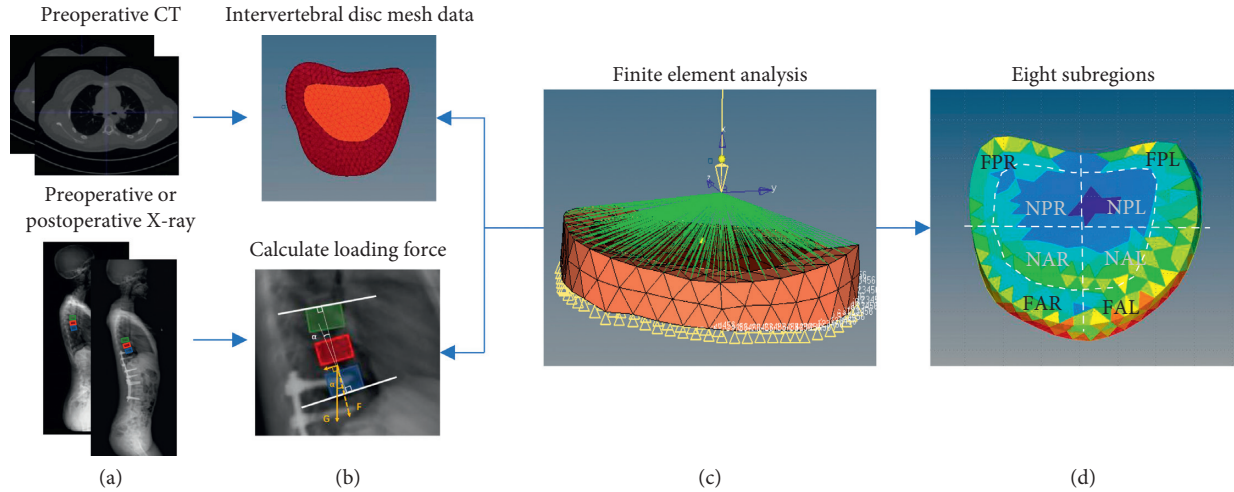


FIGURE 2: Illustration of the biomechanical analysis process. (a) Blue rectangle means the upper instrumented vertebra (UIV), the red one is UIV + 1, and the green one is UIV + 2. (b) In the mesh data of segmented cephalad intervertebral disc of UIV, light red indicates nucleus, and dark red demonstrates annulus fibrosus; white solid lines in an enlarger X-ray represent the inferior endplate of UIV and the superior endplate of two cephalad vertebrae, respectively, and the angle formed by the intersection of them is the proximal junctional angle (PJA, α). (c) Loading force calculated in the previous stage on the segmented vertebra by Altair OptiStruct. (d) F, annulus fibrosus; N, nucleus; A, anterior; P, posterior; R, right; L, left.

2.1.2. Loading Force Calculation with the PJA. To study the mechanical behavior of the proximal intervertebral disc, we calculated the loading force on the tissue before and after the surgical correction. We denote α as the PJA at pre- or postoperation. The loading force \mathbf{F} (i.e., contact force) is perpendicular to the upper surface of the intervertebral disc tissue, which is roughly equivalent to the decomposing force of the patient's gravity at this point. The decomposition angle of gravity can be estimated as α . Hence, \mathbf{F} can be calculated by $\mathbf{G} * \cos\alpha$, where \mathbf{G} denotes the gravity of the body weight above UIV (as shown in Figure 2(b)) [14].

The 3D segmented intervertebral discs were first discretized into small mesh by HyperMesh (Altair, USA). To obtain high precision mesh models, they were composed of tetrahedral elements, and each element contained 4 mesh nodes tetrahedral. The mesh nodes could be classified into the boundary and free nodes; meanwhile, the boundary nodes were located in the inferior surface, which would be fixed in all degrees of freedom. We restricted the zone to the adjacent intervertebral disc of the UIV for focusing on the deformable area of interests with regard to PJK.

2.1.3. Assignment of the Intervertebral Disc Properties. Intervertebral disc (IVD) tissue is composed of a nucleus pulposus (translucent gel) and an annulus fibrosus (lamellar structure), with negligible vascularization in the annulus and nucleus regions [23, 24]. Nonoriented collagen fibrils enmesh in the proteoglycan-water pulposus and are surrounded by the annulus fibrosus, a series of concentric encircling lamellae with two well-defined axes of orientation [23, 25–27]. To simplify the analysis process, we defined the intervertebral disc tissue as a linear elastic tissue with the homogenous and isotropic properties [5]. Different material parameters in terms of Young's modulus and Poisson's ratio

for these two components were given by the previous work depending on biomechanical analysis and the material model as shown in Table 1 [5, 28]. These parameters were used to simulate tissue biomechanical behavior based on Hooke's law. And the heterogeneous properties of the intervertebral disc tissue will be examined in the proposed studies.

2.2. Stress Formulation. We extracted stresses as one of the biomechanical characteristics from FEM. The stress for each mesh node varies according to different components of the intervertebral disc. To obtain a distribution of stress features, we first simulated the intervertebral disc behavior responding to loading force. Denote stress features as $\sigma_i = \sigma(\mathbf{G}_i, \alpha_i, \mathbf{E}, \mathbf{v})$, where G_i denotes the force of gravity from the body weight above UIV of the i th patient, and α_i is the pre- or postoperative PJA. Next, the stress value σ_i is employed as the biomechanical feature of the i th patient. Finally, $\sigma(\cdot)$ represents the stress modeled by Hooke. Here, FEM was implemented in the commercial software Altair OptiStruct. To validate our model, we loaded the pressure difference of pre- and postoperative force of one patient on his FEM model and compared the simulation results with the true geometric model of the intervertebral disc generated from postoperative CT; the absolute error of volume was 500 mm^3 , while the relative error was 2.5%, and the absolute error of average disc height on the central sagittal plane was 0.3 mm, while the relative error was 5.0%, indicating that our simplified model could save labor and machine time based on not influencing the authenticity of the FEM model [17].

To elaborate the stress variations of corresponding regions caused by the selection of PJA, stress information was evaluated on eight anatomical regions of annulus fibrosus and nucleus: left anterior, left posterior, right anterior, and right

TABLE 1: Material parameters of the intervertebral disc tissue [5, 28].

	Young's modulus (MPa)	Poisson's ratio
Nucleus	1.0	0.49
Anulus fibrosus	3.4	0.45

posterior. The difference value $\Delta\sigma$ of before and after operation subregions max or average stress information was considered as the input biomechanical features.

2.3. Model Building. Surgical outcomes, i.e., PJK are closely related to the biomechanical properties such as stress induced by curvature rectification of scoliosis [11]. A two-layer

$$\hat{W} = \underset{W}{\operatorname{argmin}} \left\| y_i - g \left[\Delta\sigma(G_i, \alpha_i, E, \nu)^1, \Delta\sigma(G_i, \alpha_i, E, \nu)^2, \dots, \Delta\sigma(G_i, \alpha_i, E, \nu)^8, \text{age}_i, \text{gender}_i, W \right] \right\|_{L2 \text{ norm}}. \quad (2)$$

The optimization is achieved by using Adam algorithm with calculating the exponentially weighted moving average of the gradient and then squaring the calculated gradient [20]. This optimal \hat{W} is fixed for PJA simulation for new patients. The risk of PJK can be predicted from this trained model. Then, we will apply a dynamic optimization method for considering other basic clinical factors once new patients are added to the prediction model.

2.4. Clinical Outcome Prediction and the Optimal Postoperative PJA. The trained model becomes $y_i = g[\Delta\sigma(G_i,$

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \left\| 0 - g \left[\Delta\sigma(G_i, \alpha_i, E, \nu)^1, \Delta\sigma(G_i, \alpha_i, E, \nu)^2, \dots, \Delta\sigma(G_i, \alpha_i, E, \nu)^8, \text{age}_i, \text{gender}_i, \hat{W} \right] \right\|_{L2 \text{ norm}}. \quad (3)$$

Thus, the optimal loading force $\hat{\alpha}$ can be estimated preoperatively to ensure a more successful operation.

To identify the most relevant features with a high degree of discrimination between PJK and non-PJK groups, we used the DX score feature selection method, whose effectiveness and efficiency have been confirmed [21]. We selected the top 5 features to perform sensitivity analysis to explore the model output variation upon a range of 5% perturbation of those important variables.

3. Results

3.1. General Information. A summary of the collected case data and the preoperative and postoperative geometric indices is provided in Table 2. Twelve cases were recruited, among which 8 were females and 4 were males with an average operation age of 16 years, ranging from 13 to 20 years and an average weight of 49 kg, ranging from 32.5 to 71 kg. The preoperative PJA in PJK and non-PJK group was $8.4^\circ \pm 2.9^\circ$ (between 5.2° and 9.4°) and $6.7^\circ \pm 5.3^\circ$ (between 0.9° and 18°),

fully connected network was employed to efficiently model nonlinear functions with less parameters [18, 19]. The clinical outcome y_i /whether PJK occurred or not of i th patient can be modeled as the following equation:

$$y_i = g \left[\Delta\sigma(G_i, \alpha_i, E, \nu)^1, \Delta\sigma(G_i, \alpha_i, E, \nu)^2, \dots, \Delta\sigma(G_i, \alpha_i, E, \nu)^8, \text{age}_i, \text{gender}_i, W \right], \quad (1)$$

where N is the total number of subregions, $g(\cdot)$ is the fully connected network, and $\Delta\sigma$ denotes the max or average difference value of stress in corresponding subregions, including 16 variables. W is the parameter (the coefficients of the variables in $g(\cdot)$) to be determined by minimizing an objective function as

$\alpha_i, E, \nu)^1, \Delta\sigma(G_i, \alpha_i, E, \nu)^2, \dots, \Delta\sigma(G_i, \alpha_i, E, \nu)^8, \text{age}_i, \text{gender}_i, \hat{W}$. When a new patient comes to the hospital, expected Y should be set to be 0, which indicates that patients will be without PJK postoperatively. \hat{W} is known (from training step), and E, ν are the subregions of the intervertebral disc tissue that can be obtained from CT data by FEM from all patients. Age at surgery and gender can be collected from their demographic data. Then, we employ Adam algorithm to preoperatively estimate the idealized optimal value of $\hat{\alpha}$ for a new individual expecting no PJK as following:

respectively, whereas the immediate postoperative (3–7 days after surgery) PJA was $13.0^\circ \pm 4.0^\circ$ (between 9.3° and 17.2°) and $8.0^\circ \pm 4.7^\circ$ (between 1.7° and 16.6°), respectively.

3.2. Model Performance. To complete the decision-making procedure prior to the surgery and reduce the medical cost, we proposed our reliable system for AIS patients. And to avoid overfitting, leave-one-out cross-validation was implemented to assess the accuracy of our approach. More specifically, one of all 12 patients was used for model testing while the rest for training, and these procedures were repeated until each patient had been used once as a testing sample. We evaluated the performance based on the difference between the predicted clinical results (PJK) and ground truth derived from 2 years of the follow-up study. The average prediction accuracy was 83.3% for all cases (2 out of 12). Incorrect predictions from the biomechanical and machine learning approach model for the patients are shown in Figure 3. The receiver operating characteristic (ROC)

TABLE 2: Optimization performance on PJK and non-PJK group.

No.	PJK after surgery	Weight (kg)	Age at operation (year)	Gender (male: M, female: F)	Predicted results (0 = without PJK and 1 = with PJK)	Pre-PJA (°)	Applied post-PJA (°)	Optimal post-PJA (°)
1	No	56	15	M	0.425	12.4	12.9	12.6
2	No	42	17	F	0.296	5.8	7.7	6.3
3	No	64	13	F	0.345	6.9	7.8	7.4
4	No	52	20	M	0.174	5.2	9.6	3.9
5	No	37	13	F	0.270	2.4	1.7	3.4
6	No	45	16	M	0.331	18.0	16.6	18.2
7	No	71	20	F	0.556 (yes)	0.9	4.8	1.2
8	No	41	13	F	0.306	4.6	3.2	6.0
9	No	40	13	F	0.361	5.1	7.6	6.1
10	Yes	32.5	16	M	0.508	10.7	17.2	11.5
11	Yes	60	18	F	0.401 (no)	4.1	12.4	5.5
12	Yes	47.5	18	F	0.586	9.4	9.3	9.5

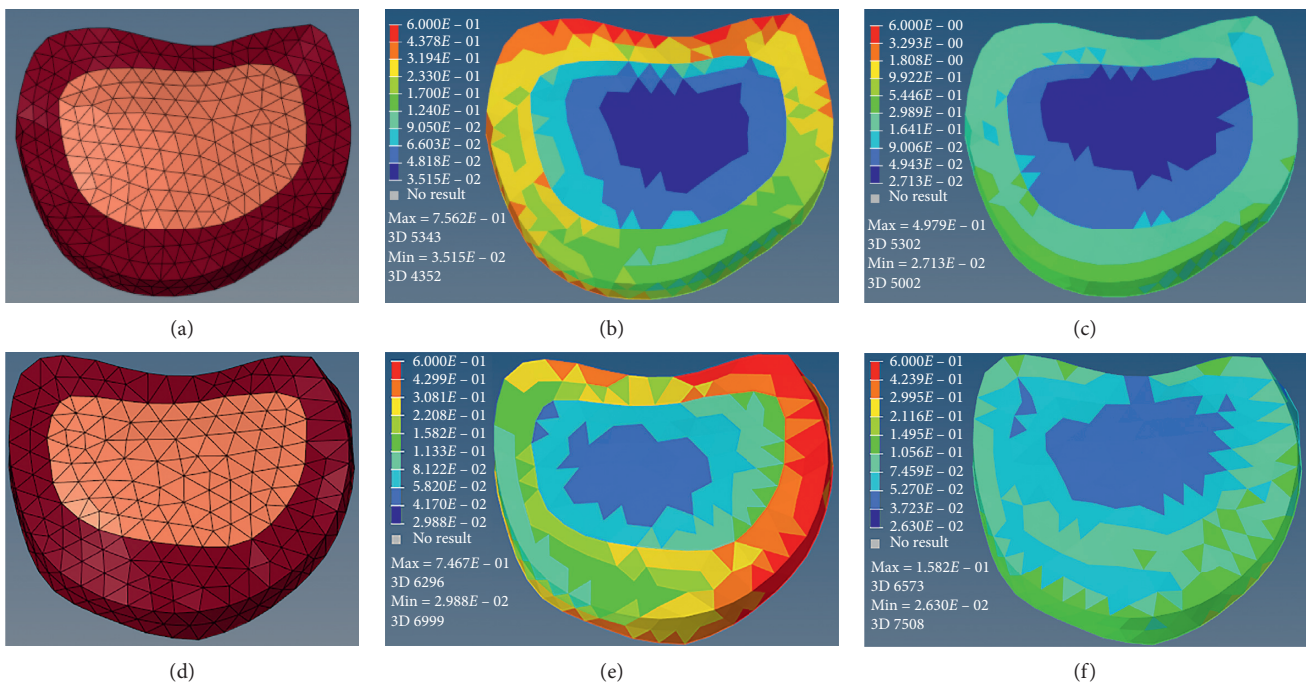


FIGURE 3: Incorrect predictions from the biomechanical and machine learning approach model for the patients in a dataset (a and d, meshed model of the intervertebral disc; b and e, preoperative stress distribution; c and f, postoperative stress). (a–c) Case 7 was predicted to be high risk (model output, 55.6%). However, PJK was not observed until a 2-year postoperative follow-up. (d–f) The model output was 40.1% for the Case 11, in which PJK occurred 451 days after operation.

curve is illustrated in Figure 4, and the area under the curve (AUC) of prediction was 0.889.

The top 5 features with 74.3% impact percentage of all features that affected the outcome were age, max-stress variations in right anterior annulus fibrous and nucleus, and average stress variations in right anterior and left posterior nucleus. Sensitivity analysis showed that our model was stable in the sense that the output variance was less than 5% when the important factor values were perturbed in a range of 5%. The top five factors ranked by the DX score were sensitive for all patients (2.10%–4.66% upon 5% parameter perturbation) in Figure 5.

4. Discussion

This study is the first to integrate computational biomechanics and machine learning to generate clinically relevant results for surgical scoliosis treatment. A procedural planning tool was constructed to focus on predicting the risk of PJK in scoliosis patients undergoing spine surgery through computation of stress within the intervertebral disc and optimization of the PJA to prevent PJK. In clinical practice, when making a surgical plan for the new patient, we can accurately simulate his or her postoperative biomechanical behavior on proximal intervertebral disc and predict the risk

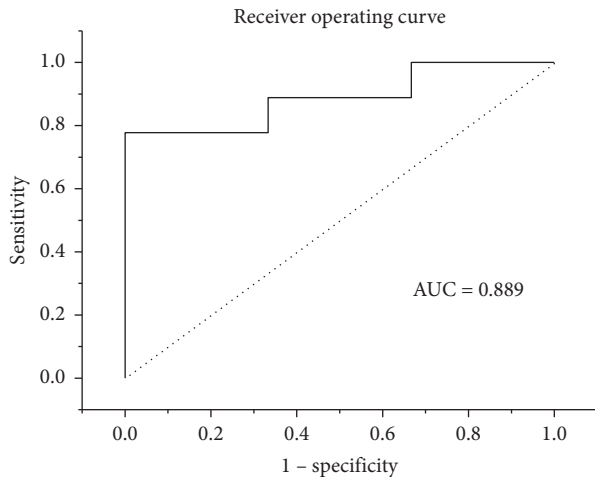


FIGURE 4: ROC curves of the two-layer fully connected network model.

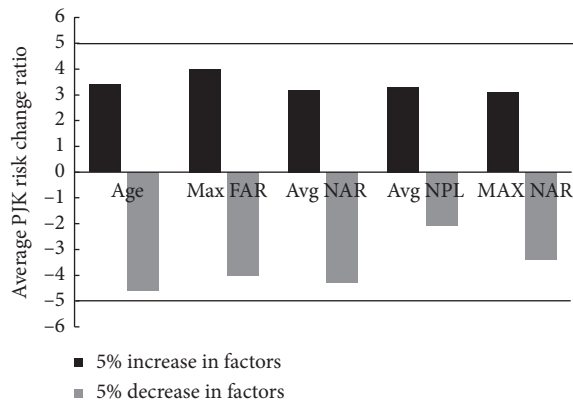


FIGURE 5: Sensitivity analysis for the top 5 features ranked by DX score: age, max-stress variations in right anterior annulus fibrosus (Max FAR) and nucleus (Max NAR), and average stress variations in right anterior (Avg NAR) and left posterior nucleus (Avg NPL).

of PJK and obtain the best optimal PJA with a learned model. Therefore, the decision-making procedure could be performed prior to the surgery, reducing the risk of revision related to postoperative complications.

Cases 1–9 underwent spinal corrective operation without PJK, and Cases 10–12 who had PJK after spinal corrective operation required being followed-up closely and received revision in time. Table 2 represents the predicted performance of our approach. The average of prediction accuracy was 83.3% for all cases (10 out of 12). Certain cases failed for reasons which could not be attributed to the model. For example, Case 7 failed because this patient received growing rod technique in our hospital before being treated by long posterior instrumentation and fusion surgery, which was not our research object in this model. The Case 8 might be due to vertebroplasty, a technique injecting bone cement into vertebral bodies to restore the stiffness and increasing the risk of fractures in adjacent nonaugmented vertebrae [29]. Further investigation is required to determine the biomechanical effect of vertebroplasty in developing PJK, and it is beyond the scope of this study. Our model indicated that some regions

with high difference of stress before operation and after operation were observed in the PJK group. These excessive stress concentrations demonstrated that inappropriate intraoperative change of PJA would lead to a nonsynchronous variation of stress in the corresponding areas of the intervertebral disc. And the structural changes of intervertebral discs might be attributed to the inhomogeneous distribution of stress on the surface, which eventually led to an increase in the PJA. Moreover, increased PJA was associated with the clinical outcome in elderly populations [30].

Spinal instrumentation offers benefits to AIS patients whose cardiopulmonary function and growth are seriously affected by the spinal deformity [31, 32]. However, complicated issues that adversely affect surgery outcomes remain, and proper operation design may help address current challenges. For spinal instrumentation surgery, to guarantee the corrected spine is aligned together in three dimensions at the macro level; it means that there will be an inevitable risk of neglecting the detail treatment of the proximal junction region, especially PJA. For these reasons, Lee et al. [9] were the first to suggest fixing UIV + 1 if it was more than 5° to reduce the incidence of PJK after operation. However, no previous study assesses a specific and applicable PJA for each individual case with scoliosis, and our study devotes to resolve this problem. It is worthwhile to mention that simplifications and approximations have been made in the intervertebral disc FEM model construction. A more precise spine geometry may be complex and certainly time-dependent such as vertebrae and tissue degeneration, which may go beyond the biomechanical domain considered in this study. However, given that the PJK evolves primarily in the early postoperative period, the primary effects of the PJA can be estimated as mechanical modifications of the intervertebral disc such as load-stress or load-stress behaviors that in turn give rise to changes of the disk anatomical structure as function of time [16]. No tissue deformations appeared in adjacent vertebrae, and the FEM model developed in this study allowed the primary effects of the PJA to be simulated and assessed, restricting the area-of-interest to the intervertebral disc that could be considered as appropriate for this study. Belytschko et al. also indicated that material properties of the annulus obtained by direct measurement underestimated the material stiffness, and based on the disc geometry, reasonable predictions of variations of disc stiffness with vertebral level could be made [25]. Thus, the influence of material parameters associated with disc levels in surgical optimization should be the future research direction. Another simplification was that pre- and post-procedural simulations were performed on the same intravertebral disc model, which could be considered as having a limited impact on the results, because the analyses focused on the relative differences instead of the absolute numerical values of the variables.

The stability and accuracy of our established machine learning model have been tested and verified by the cross-validation and comprehensive sensitivity analyses. The practically applied PJA in operation and the estimation results from our model are simultaneously shown in Table 2. The optimal PJA estimation results from our model showed

much consistence with the actual sizes in the non-PJK group. In addition, certain optimization results (e.g., Case 5, 6, and 8) also showed that it was unreasonable to minimize PJA blindly, and the best angle should be chosen by comprehensively considering the influence of biomechanical and demographic data. In the PJK group, the estimation results could reduce the excessive stress of the proximal intervertebral disc tissue, which might delay the early deterioration and dysfunction of proximal junctional region and lead to a favorable outcome. The estimation of Case 12 was little different compared with practically adopted PJA, and it might be caused by other nonbiomechanical problems, which were not collected and were potentially better predictors for that patient.

On account of the complexity and variability of the proximal junctional anatomy, the incidence and severity of postprocedural PJK are difficult to predict, indicating the need of a model that aid the orthopedic surgeon to select the optimal PJA that best fits the individual patient. Accurate simulation of a spinal surgery procedure based upon the integration of age and gender, the patient-specific anatomy, the biomechanical properties of the intervertebral disc tissue may serve this goal. The combination of biomechanical properties and the machine learning method substantially improved prediction of clinical results. A nonlinear FEM approach will be used to improve the accuracy in the future.

Some limitations need to be considered in this study. First, we developed finite-element models with simple linear elastic material properties for AIS patients. The anisotropy properties of the intervertebral disc tissue will be examined in the proposed studies. Second, there were very small number of patients and lacking of factors other than biomechanical and basic data. However, the current study is the first step to focus on integrating computational biomechanics and machine learning to optimize the PJA for scoliosis patients, and we will keep following up more cases in next stage.

5. Conclusions

In this work, we developed a procedural planning tool to predict the risk of PJK and the optimal PJA. 3D FEM models of the cephalad intervertebral disc of UIV were constructed for all patients to extract biomechanical stress information. The two-layer fully connected network was used to model the relationships between the stress information and the risk of PJK. We have integrated biomechanics and machine learning to propose a systematic approach, which devotes to support surgical decision-making by reducing the medical cost of revision after long posterior instrumentation and fusion surgery in the future.

Data Availability

The raw data used to support this study are not available and belong to West China Hospital of Sichuan University.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank 1-3-5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (Grant ZYJC18010), 1-3-5 Project for Disciplines of Excellence–Clinical Research Incubation Project, West China Hospital, Sichuan University (Grant 2019HXFH022), and Post-Doctor Research Project, West China Hospital, Sichuan University (Grant 2019HXBH039) for their support in this study.

References

- [1] T. Mimura, J. Takahashi, S. Ikegami et al., “Can surgery for adolescent idiopathic scoliosis of less than 50 degrees of main thoracic curve achieve good results?” *Journal of Orthopaedic Science*, vol. 335, 2017.
- [2] K. H. Bridwell, “Spinal instrumentation in the management of adolescent scoliosis,” *Clinical Orthopaedics and Related Research*, vol. 335, no. 335, pp. 64–72, 1997.
- [3] T. Humke, D. Grob, H. Scheier, and H. Siegrist, “Cotrel-dubouset and harrington instrumentation in idiopathic scoliosis: a comparison of long-term results,” *European Spine Journal*, vol. 4, no. 5, pp. 280–283, 1995.
- [4] A. N. Larson, C.-E. Aubin, D. W. Polly Jr. et al., “Are more screws better? A systematic review of anchor density and curve correction in adolescent idiopathic scoliosis,” *Spine Deformity*, vol. 1, no. 4, pp. 237–247, 2013.
- [5] S. K. Ha, “Finite element modeling of multi-level cervical spinal segments (C3–C6) and biomechanical analysis of an elastomer-type prosthetic disc,” *Medical Engineering & Physics*, vol. 28, no. 6, pp. 534–541, 2006.
- [6] J. Zhao, M. Yang, Y. Yang, Z. Chen, and M. Li, “Proximal junctional kyphosis following correction surgery in the Lenke 5 adolescent idiopathic scoliosis patient,” *Journal of Orthopaedic Science*, vol. 23, no. 5, pp. 744–749, 2018.
- [7] Z. Sun, G. Qiu, Y. Zhao et al., “Risk factors of proximal junctional angle increase after selective posterior thoracolumbar/lumbar fusion in patients with adolescent idiopathic scoliosis,” *European Spine Journal*, vol. 24, no. 2, pp. 290–297, 2015.
- [8] R. C. Glattes, K. H. Bridwell, L. G. Lenke et al., “Proximal junctional kyphosis in adult spinal deformity following long instrumented posterior spinal fusion: incidence, outcomes, and risk factor analysis,” *Spine (Phila Pa 1976)*, vol. 30, no. 14, pp. 1643–1649, 2005.
- [9] G. A. Lee, R. R. Betz, D. H. Clements 3rd. et al., “Proximal kyphosis after posterior spinal fusion in patients with idiopathic scoliosis,” *Spine (Phila Pa 1976)*, vol. 24, no. 8, pp. 795–799, 1999.
- [10] A. Sebaaly, C. Sylvestre, Y. El Quehtani et al., “Incidence and risk factors for proximal junctional kyphosis,” *Clinical Spine Surgery*, vol. 31, no. 3, pp. E178–e183, 2018.
- [11] P. G. Passias, S. R. Horn, G. W. Poorman et al., “Clinical and radiographic presentation and treatment of patients with cervical deformity secondary to thoracolumbar proximal junctional kyphosis are distinct despite achieving similar outcomes: analysis of 123 prospective CD cases,” *Journal of Clinical Neuroscience*, vol. 56, pp. 121–126, 2018.
- [12] Y. J. Kim, K. H. Bridwell, L. G. Lenke et al., “Proximal junctional kyphosis in adult spinal deformity after segmental posterior spinal instrumentation and fusion: minimum five-year follow-up,” *Spine (Phila Pa 1976)*, vol. 33, no. 20, pp. 2179–2184, 2008.

- [13] S. Mendoza-Lattes, Z. Ries, Y. Gao et al., "Proximal junctional kyphosis in adult reconstructive spine surgery results from incomplete restoration of the lumbar lordosis relative to the magnitude of the thoracic kyphosis," *Iowa Orthopedic Journal*, vol. 31, pp. 199–206, 2011.
- [14] M. Yagi, K. B. Akilah, and O. Boachie-Adjei, "Incidence, risk factors and classification of proximal junctional kyphosis: surgical outcomes review of adult idiopathic scoliosis," *Spine (Phila Pa 1976)*, vol. 36, no. 1, pp. E60–E68, 2011.
- [15] W. W. Schairer, A. Carrer, V. Deviren et al., "Hospital readmission after spine fusion for AdultSpinal deformity," *Spine (Phila Pa 1976)*, vol. 38, no. 19, pp. 1681–1689, 2013.
- [16] C.-E. Aubin, M. Cammarata, X. Wang, and J.-M. Mac-Thiong, "Instrumentation strategies to reduce the risks of proximal junctional kyphosis in adult scoliosis: a detailed biomechanical analysis," *Spine Deformity*, vol. 3, no. 3, pp. 211–218, 2015.
- [17] S. Bess, J. E. Harris, A. W. L. Turner et al., "The effect of posterior polyester tethers on the biomechanics of proximal junctional kyphosis: a finite element analysis," *Journal of Neurosurgery: Spine*, vol. 26, no. 1, pp. 125–133, 2017.
- [18] M. Cammarata, C. E. Aubin, X. Wang et al., "Biomechanical risk factors for proximal junctional kyphosis: a detailed numerical analysis of surgical instrumentation variables," *Spine (Phila Pa 1976)*, vol. 39, no. 8, pp. E500–E507, 2014.
- [19] P. J. Cahill, W. Wang, J. Asghar et al., "The use of a transition rod may prevent proximal junctional kyphosis in the thoracic spine after scoliosis surgery: a finite element analysis," *Spine (Phila Pa 1976)*, vol. 37, no. 12, pp. E687–E695, 2012.
- [20] L. Fradet, X. Wang, L. G. Lenke, and C.-E. Aubin, "Biomechanical analysis of proximal junctional failure following adult spinal instrumentation using a comprehensive hybrid modeling approach," *Clinical Biomechanics*, vol. 39, pp. 122–128, 2016.
- [21] S. Ghailane, S. Pesenti, E. Peltier, E. Choufani, B. Blondel, and J. L. Jouve, "Posterior elements disruption with hybrid constructs in AIS patients: is there an impact on proximal junctional kyphosis?" *Archives of Orthopaedic and Trauma Surgery*, vol. 137, no. 5, pp. 631–635, 2017.
- [22] D. J. Pearsall, J. G. Reid, and L. A. Livingston, "Segmental inertial parameters of the human trunk as determined from computed tomography," *Annals of Biomedical Engineering*, vol. 24, no. 2, pp. 198–210, 1996.
- [23] A. Gloria, F. Causa, R. De Santis, P. A. Netti, and L. Ambrosio, "Dynamic-mechanical properties of a novel composite intervertebral disc prosthesis," *Journal of Materials Science: Materials in Medicine*, vol. 18, no. 11, pp. 2159–2165, 2007.
- [24] A. Malandrino, J. Noailly, and D. Lacroix, "The effect of sustained compression on oxygen metabolic transport in the intervertebral disc decreases with degenerative changes," *PLoS Computational Biology*, vol. 7, no. 8, 2011.
- [25] T. Belytschko, R. F. Kulak, A. B. Schultz, and J. O. Galante, "Finite element stress analysis of an intervertebral disc," *Journal of Biomechanics*, vol. 7, no. 3, pp. 277–285, 1974.
- [26] A. Gloria, T. Russo, U. D'Amora, M. Santin, R. De Santis, and L. Ambrosio, "Customised multiphasic nucleus/annulus scaffold for intervertebral disc repair/regeneration," *Connective Tissue Research*, vol. 61, no. 2, pp. 152–162, 2020.
- [27] T. Belytschko and R. F. Kulak, "A finite-element method for a solid enclosing an inviscid, incompressible fluid," *Journal of Applied Mechanics*, vol. 40, no. 2, p. 609, 1973.
- [28] H. W. Ng, E. C. Teo, and V. S. Lee, "Statistical factorial analysis on the material property sensitivity of the mechanical responses of the C4-C6 under compression, anterior and posterior shear," *Journal of Biomechanics*, vol. 37, no. 5, pp. 771–777, 2004.
- [29] U. Berlemann, S. J. Ferguson, L.-P. Nolte, and P. F. Heini, "Adjacent vertebral failure after vertebroplasty," *The Journal of Bone and Joint Surgery. British Volume*, vol. 84-B, no. 5, p. 748, 2002.
- [30] T. Raman, E. Miller, C. T. Martin et al., "The effect of prophylactic vertebroplasty on the incidence of proximal junctional kyphosis and proximal junctional failure following posterior spinal fusion in adult spinal deformity: a 5 year follow up study," *Spine Journal*, vol. 31, 2017.
- [31] M. B. Dobbs, L. G. Lenke, Y. J. Kim et al., "Anterior/posterior spinal instrumentation versus posterior instrumentation alone for the treatment of adolescent idiopathic scoliotic curves more than 90°," *Spine (Phila Pa 1976)*, vol. 31, no. 20, pp. 2386–2391, 2006.
- [32] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *Computer Science*, vol. 31, 2014.