

Research Article

Nanotechnology-Based Sensitive Biosensors for COVID-19 Prediction Using Fuzzy Logic Control

Vikas Maheshwari,¹ Md Rashid Mahmood,¹ Sumukham Sravanthi,² N. Arivazhagan,³ A. ParimalaGandhi,⁴ K. Srihari ,⁵ R. Sagayaraj,⁶ E. Udayakumar,⁴ Yuvaraj Natarajan,⁷ Prashant Bachanna,⁸ and Venkatesa Prabhu Sundramurthy ⁹

¹Dept. of ECE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Dist R.R., Hyderabad Telangana State 501506, India

²Dept. of CSE, Kakatiya Institute of Technology and Science, Warangal, India

³Department of Computational Intelligence, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur 603203, India

⁴Dept. of ECE, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, India

⁵CSE, SNSCT, India

⁶Department of EEE, Muthayammal Engineering College, Namakkal, India

⁷Research and Development, ICT Academy, 600096, Chennai, India

⁸Dept. of ECE, Bharat Institute of Engineering and Technology, Hyderabad, Telangan 501510, India

⁹Department of Chemical Engineering, Addis Ababa Science and Technology University, Ethiopia

Correspondence should be addressed to Venkatesa Prabhu Sundramurthy; venkatesa.prabhu@aastu.edu.et

Received 4 September 2021; Revised 22 September 2021; Accepted 7 October 2021; Published 3 November 2021

Academic Editor: Lakshmiopathy R

Copyright © 2021 Vikas Maheshwari 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.

Increasing the growth of big data, particularly in healthcare-Internet of Things (IoT) and biomedical classes, tends to help patients by identifying the disease early through methods for the analysis of medical data. Hence, nanotechnology-based IOT biosensors play a significant role in the medical field. *Problem.* However, the consistency continues to decrease where missing data occurs in such medical data from nanotechnology-based IOT biosensors. Furthermore, each region has its own special features, which further lowers the accuracy of prediction. The proposed model initially reconstructs lost or partial data in order to address the challenge of handling the medical data structures with incomplete data. *Methods.* An adaptive architecture is proposed to enhance the computing capabilities to predict the disease automatically. The medical databases are managed by unpredictable environments. This optimized paradigm for diagnosis produces the fuzzy, genetically categorized decision tree algorithm. This work uses a normalized classifier namely fuzzy-based decision tree (FDT) algorithm for classifying the data collected via nanotechnology-based IOT biosensors, and this helps in the identification of nondeterministic instances from unstructured datasets relating to the medical diagnosis. The FDT algorithm is further enhanced by using genetic algorithms for effective classification of instances. Finally, the proposed system uses two larger datasets to verify the predictive precision. In order to describe a fuzzy decision tree algorithm based upon the fitness function value, a modified decision classification rule is used. The structure and unstructured databases are configured for processing. *Results and Conclusions.* This evaluation of test patterns helps to track the efficiency of FDT with optimized rules during the training and testing stages. The proposed method is validated against nanotechnology-based IOT biosensors data in terms of accuracy, sensitivity, specificity, and *F*-measure. The results of the simulation show that the proposed method achieves a higher rate of accuracy than the other methods. Other metrics relating to the model with and without feature selection show an improved sensitivity, specificity, and *F*-measure rate than the existing methods.

1. Introduction

As medical knowledge grows, the electronic health record (EHR) subsequently grows dramatically. The COVID-19 prediction has become a major factor in big data analytics as data increases in size. Classification algorithms were developed to improve medical diagnosis accuracy [1]. The classification method in big data analysis leads to the classification of datasets according to the diagnostic application with machine learning algorithms.

Efficient techniques are employed in large data analytics to find insights, correlations, and cached patterns from input data collected from the nanotechnology-based IoT biosensors. The data analytics provides improved decision-making, cost reductions, and development of new items to meet the customer requirements. Hence, it addresses the challenges of various applications, such as health care, plants, and bioinformatics, with wide advantages [2]. The problems are addressed via machine learning strategies that include rule-based and decision-making. Most classification algorithms only take structured data into account. In the processing of unstructured data, structured and unstructured information is generally combined [3, 4] to reduce the disease-prediction risk. The combination of the information eases the cost of processing and reduces the redundant information.

Artificial Intelligence (AI) is a troublesome technology used as smarter technology on wide varied applications, ranging from automobile industry to healthcare industry. AI was additionally used to track the virus spread, to identify the patients with high risk areas and antiviral drug in controlling the pandemic in real-time environment. AI predicts the risk of mortality rate by analyzing the patient datasets. This application of AI can help in screening the population, notification, medical help, and suggestion on infection control. It further assists in treatment, planning, and prediction of disease spread and outcomes of patients using AI evidence-based tool.

AI is a powerful intelligence method in the fighting the pandemic, and it has scrambled the AI on healthcare analytics. AI with predefined datasets can predict and track the infectious spread on timely manner across various regions. The challenges include problems associated with forecasting the pandemic over unbiased and historical data for training the AI. It includes panic activities among humans and the statistics difference from existing pandemics (Spanish flu, H1N1 influenza, and AIDS). The lack of proper datasets and big data is considered problematic in finding the infectious spread.

Therefore, the classification of unstructured data using classification algorithms is important to classify separately. The risk of disease prediction based on qualified classifiers is reduced in this manner. Structured data treatment method is proposed in the [5] for unstructured medical image data. Integrated structure systems [6] for medical text documents are structured using a Bayesian classifier for extracting the attributes. In addition, k -means identifies the data and ensures optimum data classification. The search method [7, 8] is used to classify the connections through which unstructured medical data is organized. This proposed technique

TABLE 1: List of AI-based classification.

Classification model	Dataset	Accuracy
3D CNN model	498 CT scans	70.02
Desenet201	1260 images	96.21
3D CNN	413 CT images	93.01
ResNet-50	60,457 CT images	98.81
RF and SVM	626 CT images	83.77
DL model	219 images	86.72

produces improved accuracy results than the other techniques based on the SVM [9–13]. Some of the methods are listed in Table 1.

The above-mentioned methodology cannot identify medical datasets through a rule-based system. The processing of datasets from the input nanotechnology IoT sensitive biosensors [14–16] was done using a rule-based method to minimize redundant data. The rule-based framework with its rule base and systems unregulated data removes redundant data. In order to improve the risk of accuracy in classification, the rule set is needed.

In this paper, we propose a fuzzy decision tree (FDT) method for classification, thereby enhancing the novelty using the genetic algorithm (GA) to improve decision-making on a rule-based basis in broad unstructured datasets from nanotechnology-based IoT biosensors.

The main contribution of the paper is as follows:

- (i) The work uses a normalized classifier namely the fuzzy-based decision tree (FDT) algorithm to identify the nondetermined instances relating to the medical diagnosis due to the unstructured nature of the datasets from nanotechnology-based IoT biosensors
- (ii) The genetic algorithm(GA) is used to improve the FDT algorithm's classification rule collection
- (iii) The evaluation of test patterns helps to track the efficiency of FDT with optimized rules during the training and testing stages
- (iv) Finally, the proposed system uses two larger datasets to verify the likelihood of predictive precision.

In this proposed work, Section 2 provides the concept of the article. Section 3 discusses the FDT with GA to design the predictive problem. Section 4 validates the entire work. Section 5 concludes the paper with possible directions of future scope.

2. Basic Concept

This section provides the basics of the hesitant fuzzy algorithm (HFA) that eliminates the hesitations associated with fuzzy set assignment and membership degree to process the data from nanotechnology-based IoT biosensors. The following provides the HFA preliminaries:

For a fuzzy set D with reference set (Y) is generally represented in terms of a function $hD(y)$, where Y produces a subset $[0, 1]$, when D is applied over the Y :

$$D = \{ \langle y, h(y) \rangle | y \in Y \} \text{ where } y \in Y, \quad (1)$$

where the membership degree of $y \rightarrow A$ is defined as different values set $hD(y)$. Hence, for simplicity, the $h(y)$ is generally referred to as a fuzzy set element.

3. Fuzzy Decision Tree (FDT)

In this section, we provide FDT details and how the genetic algorithm provides the optimized rule for FDT as illustrated in Figure 1.

3.1. Data Balancing. Unbalanced datasets collected from nanotechnology-based IoT biosensors. The former model tends to reduce the high-dimensional samples and do not take useful information from the account. Samples for the small class can be oversized by the latter procedure. At first, k -means collects the samples from several classes in various clusters. A number of pseudoclasses were marked or numerated for the classes of balanced dataset collected from nanotechnology-based IoT biosensors.

3.2. Construction of FDT. Instances with differential membership are permitted to use FDT from $[0, 1]$ to multiple branches. Using fuzzy rules, the node conditions of a branch are specified. The cases are degraded by different membership levels as they fall at different nodes. If the information or noise is incomplete, the downfall of the cases is considered beneficial. However, FDT is slow to use, but the ranking is better than an ordinary tree.

The FDT consists of construction of tree and nodes for optimal decision-making. FDT is a fuzzy logic algorithm, which uses language terms to change the attributes of the data on medical training. The knowledge gained is used for attribute evaluation on the connected node. It also uses a fuzzy dataset that includes membership, input, and target attributes. The child node set includes all instances of parent nodes that delete branch attributes. Furthermore, in all cases, the main distinction occurs in the fuzzy membership.

Consider an input preprocessed dataset (S) collected from collected from nanotechnology-based IoT biosensors with an attribute (A_i), where the study uses $\{F_i^{(1)}, \dots, F_i^{(r_i)}\}$. The membership degree of is given by

$$\mu_S[F_i^{(j)}](X_e) = \mu_{F_i^{(j)}}(x_e^{(i)}) \times \mu_S(X_e), \quad (2)$$

where $\mu_S(X_e)$ is defined as the membership degree of X_e , $\mu_{F_i^{(j)}}(x_e^{(i)})$ is defined as the membership degree of $x_e^{(i)}$ with a fuzzy term $F_i^{(j)}$, and $S[F_i^{(j)}]$ is defined as the child node.

The algorithm takes the branch attribute into consideration, based on Figure 2, based on the maximum data gain fuzzy value indicating the fuzzy entropy.

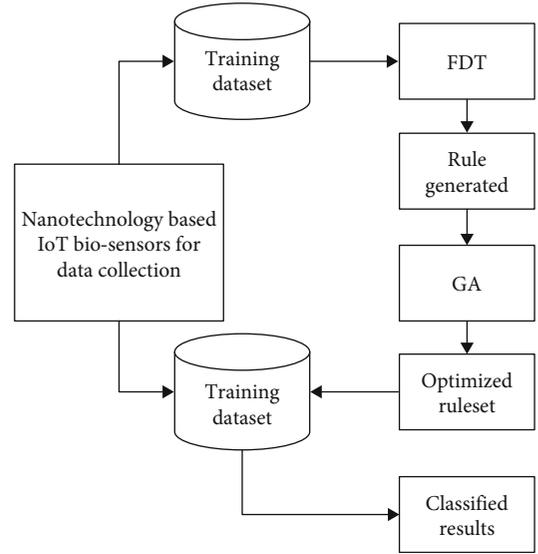


FIGURE 1: Proposed classification framework.

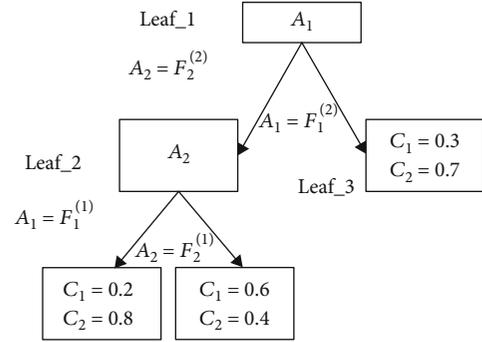


FIGURE 2: Fuzzy decision tree.

$$E(S) = \sum_{i=1}^m - \frac{|S_{y=c_i}|}{|S|} \log_2 \frac{|S_{y=c_i}|}{|S|}. \quad (3)$$

Information gain is given by

$$I_G(A_i, S) = E(S) - \sum_{j=1}^{r_i} w_j E(S[F_i^{(j)}]), \quad (4)$$

where $E(S)$ is defined as the entropy function of S , $E(S[F_i^{(j)}])$ is defined as the child node (j) entropy, and w_j is defined as the child node (j) instances and it is given as follows:

$$w_j = \frac{|S[F_i^{(j)}]|}{\sum_{k=1}^{r_i} |S[F_i^{(k)}]|}. \quad (5)$$

The above algorithm illustrates the FDT implementation process. The implementation of the FDT uses a stopping criterion. The standardised maximum I_G [28] method is used as a stopover criterion.

Inputs: membership function, training data, threshold value.
 Membership function is set as unity.
 Generate root node using fuzzy set.
 For a node (N).
 Check if the end criteria is reached,
 Assume the existence of a leaf N .
 Mark the record N belongs to a class is labelled.
 Else if end criteria is not reached, then,
 Estimate the I_G .
 Estimate maximum I_G .
 Find child nodes.
 End.
 End

ALGORITHM 1: FDT algorithm.

The FDT builds the decision tree using discrete procedure in which a fuzzy system is specified for a certain attribute A_i :

$$B = \{ \langle A_i, e_B(A_i) \rangle | A_i \in A \}. \quad (6)$$

Finally, the HI_G over an attribute A_i on a dataset (S) is given by

$$HI_G(A_i, S) = 0.2 \left(I_G \left(f_{uf}(A_i, S) \right)^2 + I_G \left(f_{in}(A_i, S) \right)^2 \right). \quad (7)$$

Here, Equation (7) is computed using two energy values: (i) merging and (ii) uniform discretization.

$$e_B(A_i) = \left\{ I_G \left(f_{uf}(A_i, S) \right), I_G \left(f_{in}(A_i, S) \right) \right\} \quad (8)$$

$F(A_i, S)$ is defined as the predicted drop of entropy due to an attribute A_i .

$f_{in}(A_i, S)$ is defined as the predicted drop in entropy (discretization merge) due to an attribute A_i .

$f_{uf}(A_i, S)$ is defined as the predicted drop in entropy due to an attribute A_i (uniform frequency).

The generated intervals of parameter n are considered the same and represent a controlled process. In this node, the fuzzy discretization method based on the information gain is selected using discrete methods.

3.3. FDT Inference Engine. The decision tree is considered a rule in the form of leaves. This condition contributes to the combined history and is classified as a leaf. The rule set is regarded as consistent when a single classification is performed between the leaves. The main significant for a rule is also known as consistent training information and a set of appropriate characteristics. However, when the fuzzy set is inconsistent, the nonnull membership function results from its fuzzy value over a single fuzzy set.

The FDT is then transformed into a fuzzy decision trees in Figure 2. The fuzzy rule fits every leaf. The approximate reasoning operates under four different categories that

include (i) firing strength, (ii) compatibility degree, (iii) certainty degree, and (iv) overall output.

3.4. Membership Function Generation. In this paper, two discretizing methods are used for the cutting points and triangular membership functions. The SD membership feature transforms the left/right functions into trapezoidal ones, where the left and right median values are the same. Finally, both discretization functions generate and build up a membership function.

4. Rule Optimization Using GA

The decision-making with fuzzy cannot change the membership rules in order to obtain maximal results. As a result, the genetic algorithm is used as a primary factor to optimize rules using FDT output optimization.

The FDT will be used first in this article to generate classification rules. The GA builds the fitness function based on the advantages of classification and precision. The rule will be optimized if the genetic value is greater and vice versa. To optimize fitness with the crossover and mutation function, the fitness function is modified. With this change, the rules get simplified.

4.1. Coding for the Rule. GA uses the binary code, which is a fixed-long bitcode used in strings with the $\{0, 1\}$ symbol as the human symbol. The encoding length is determined by the attributes value which affects the various GA bitcodes. If k -means values in an attribute, they are distributed into k bits having a value. The chromosomes are long and easier to convert in GA.

The key disadvantage of the FDT tree is the absence of discrete and numerical characteristics. A number of secondary steps must therefore be taken in conjunction with the binary code. The chromosome is set by the law of classification of instances. As the problem can be resolved by some chromosomes, its consistency would determine the rule set. If the rule set recognises a new sample, the GA selects the best rule, and GA selects next rule if the rules set does not recognise a new sample. Hence, if the rules do not recognise the new model, then the GA classifies this instance as a

default situation. There are a number of regulations based on its genetic priority; the chromosome is competitive with other chromosomes.

The proposed model provides the details of the genes in terms of fixed length and chromosomes:

- (i) *Weight*. Boolean variable of an attribute
- (ii) *Value*. Attribute value: continuous or discrete
- (iii) *Operator*. Genes conjunction: continuous or discrete
- (iv) *Gain ratio*. Information gain (I_G).

The rules are thus obtained with the four characteristic with a fixed and variable length of chromosome.

4.2. FDT Optimization Using GA. When rules are created by FDT over a subset with few rules on attaining a higher classification rate, the optimization process is applied. The aim is, therefore, to improve the accuracy of FDT by reducing the fitness function. The fitness function at GA tests the rule consistency. The fitness function is divided into four classes based on the rules established:

- (i) Class_A predicts true value as true and false value as true
- (ii) Class_B predicts true value as true and false value as fault
- (iii) Class_C predicts true value as fault and false value as true
- (iv) Class_D predicts true value as fault and false value as fault.

Therefore, the accuracy defines the fitness function:

$$\text{Accuracy or } A = \frac{\text{Class_A} + \text{Class_D}}{T + F}, \quad (9)$$

where T is defined as the true data sample and F is defined as the fault data sample.

The precision is capable of producing the correct results for classification:

$$\text{Support or } S = \frac{\text{Class_A} + \text{Class_C}}{T + F}. \quad (10)$$

The rule is larger if the value of the support is higher in dataset and the fitness is estimated as

$$\text{fitness or } f = \frac{N_{att} + n_{r-att}}{N_{att}}, \quad (11)$$

where N_{att} is defined as the total attributes and n_{r-att} is defined as the total attributes in a rule.

The rule is easy to understand if the individual fitness is high. Finally, it is calculated by the maximum function as well:

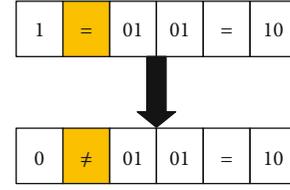


FIGURE 3: Operator mutation.

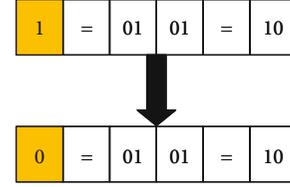


FIGURE 4: Weight mutation.

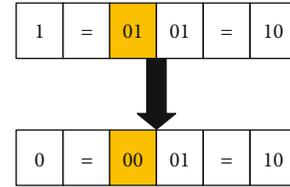


FIGURE 5: Value mutation.

$$\text{fitness}_{\max} = x \times f + y \times s + w \times a + z \times I_G, \quad (12)$$

where x , y , w , and z are defined as the variable weights lying in $[0, 1]$ and

$$x + y + w + z = 1. \quad (13)$$

4.2.1. Crossover and Mutation Operations. The sample dataset collected from nanotechnology-based IoT biosensors is coded using code rules, creating new successful individuals. Search space will then be reduced significantly, and processing speed will be increased.

This paper uses a 2-point crossover and generates an interval of random numbers $[0, 1]$. The parents are selected randomly for crossover operation if the chance of a crossover exceeds the random number. In the same way, a random number is generated by an interval $[0, 1]$, resulting in a mutation that exceeds a random number.

The genes consist of four components that must be effectively designed. It is a three-way transmission and one operator benefit ratio, which is shown in Figures 3–5, which is as follows:

- (i) *Operator mutation*. If the original gene attributes are changed, the gene will mutate into the gene and vice versa
- (ii) *Weight mutation*. Count the weight of the new gene to be zero, and vice versa. The gene attribute does not occur in this article if the weight changes from one to zero

TABLE 2: Evaluation of accuracy with and without feature selection (FS).

Algorithms	With FS	Without FS
Proposed	0.99543	0.982747
SVM-AIRF	0.995378	0.981016
SVM-GA	0.996133	0.981992
SVM-Fuzzy	0.996146	0.96138
BC	0.996901	0.963646
LDA	0.996901	0.96293
LR	0.981094	0.966107
DMDA	0.98082	0.964219
NLP	0.982174	0.965951

TABLE 3: Evaluation of F -measure.

Algorithms	With FS	Without FS
Proposed	0.881778	0.822648
SVM-AIRF	0.880592	0.802064
SVM-GA	0.901165	0.819711
SVM-Fuzzy	0.90153	0.604428
BC	0.915121	0.625737
LDA	0.915121	0.622364
LR	0.796866	0.645609
DMDA	0.795615	0.630347
NLP	0.812901	0.647051

- (iii) *Value mutation*. In the case of a different attribute, the initial gene replaces the gene value. Likewise, when a decimal value is created at random, the decimal value will be changed to + or - in the event of continuing attributes.

5. Experimental Results and Discussion

This section presents the validation of FDT used to collect data from nanotechnology-based IoT biosensors, where the sensors are supplied with the data from cord-19 dataset (available at <https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>). Performance metrics of the proposed method are evaluated as accuracy, sensitivity, specificity, and F -measure. In other classifiers, such as SVM or Bayes, the performance is measured. Parameters are chosen based on the fuzzy model setup. It means that the algorithm, while learning, optimizes these coefficients (according to a given optimization strategy) and returns an array of parameters which reduces the error.

The proposed model is evaluated using an Artificial Immune Recognition System based on support vector machine, SVM-genetic algorithm, SVM-Fuzzy, Latent Dirichlet allocation, Bayesian classifier, dictionary-based linguistic rules decision models, and natural language processing, as shown in Tables 2–7.

The results show that the FDT (Table 8) is more accurate than conventional methods. The principal reason for the

TABLE 4: Evaluation of precision.

Algorithms	With FS	Without FS
Proposed	0.826911	0.998376
SVM-AIRF	0.825347	0.998985
SVM-GA	0.869621	0.996198
SVM-Fuzzy	0.869147	0.544712
BC	0.903521	0.568991
LDA	0.903521	0.569417
LR	0.972014	0.58184
DMDA	0.986919	0.574265
NLP	0.986401	0.589668

TABLE 5: Evaluation of sensitivity.

Algorithms	With FS	Without FS
Proposed	0.944444	0.699522
SVM-AIRF	0.943764	0.669993
SVM-GA	0.935083	0.696346
SVM-Fuzzy	0.93642	0.67885
BC	0.927023	0.695057
LDA	0.927023	0.686166
LR	0.675202	0.725076
DMDA	0.666434	0.698569
NLP	0.691306	0.716806

TABLE 6: Evaluation of specificity.

Algorithms	With FS	Without FS
Proposed	0.996367	0.999931
SVM-AIRF	0.996327	0.999959
SVM-GA	0.997306	0.999834
SVM-Fuzzy	0.997293	0.974218
BC	0.998183	0.975927
LDA	0.998183	0.975825
LR	0.99887	0.976826
DMDA	0.999476	0.97635
NLP	0.999434	0.977293

TABLE 7: Validation of infections developing in a patient w.r.t risk factors (RF).

Patient	Unchangeable RF	Changeable RF	Controllable RF	Likelihood of infection
1	0.6438	0.6071	0.5946	0.6124
2	0.3241	0.4017	0.3339	0.3439
3	0.5096	0.6256	0.6842	0.6399
4	0.4389	0.3392	0.5057	0.451
5	0.696	0.6609	0.7347	0.6908
6	0.6408	0.6396	0.6725	0.6513

TABLE 8: Notations.

$\{A_1, \dots, A_k, Y\}$	Dataset attributes
Y	Target attribute
k	Total attributes
A_i	i^{th} instance of an attribute A
$\{A_1, \dots, A_k\}$	Set of attributes
y_i	Value of Y
m	Total number of classes
$Y \in \{c_1, \dots, c_m\}$	Class labels
$\mu_s(X_i)$	Membership degree
$S = \{(X_1, \mu_s(X_1)), \dots, (X_n, \mu_s(X_n))\}$	Fuzzy input dataset structure
n	Total number of instances
$\{F_i^{(1)}, \dots, F_i^{(r_i)}\}$	Fuzzy term
r_i	Total fuzzy terms
$[x_i^{(1)}, \dots, x_i^{(k)}, y_i]^T$	Attribute values
$ S = \sum_{i=1}^n \mu_s(X_i)$	Total instances
$\mu_{F_i}(j)$	Fuzzy membership function
$F_i^{(j)}$	Fuzzy term
$S[F_i^{(j)}]$	Child node
$S_{y=c_i}$	Instance

improvement is that the genetic algorithm is present to optimize the rules for structuring the large dataset collected from nanotechnology-based IoT biosensors. In the testing and training phase, the proposed method manages incomplete data. The FDT-classifier effectively manages the missing data during the preprocessing operation, which produces better results and outweighs conventional classifiers. This method also reduces the disparity in classification decisions based on their respective decision-making treaties. A new record is effectively classified on the basis of HDFT. In prediction, the presence of the feature selection method is significant in comparison with the standard methods. As the information gain is selected by its entropy values for the corresponding data, consequently, the results are increased and correctly diagnosed.

6. Conclusion

In this paper, we propose technique to improve the risk prediction for COVID-19 new classification technique incorporating FDT in genetic algorithms for rule optimization. The proposed solution is much more likely to be diagnosed by physicians compared to traditional prediction algorithms. Moreover, by using metaheuristic methods, the proposed work can be strengthened to refine the rule set of the decision tree.

Future methods can rely on finding the confirmation of real-time data over polymerase chain reaction of a viral

agent. AI-based ML/DL/RIL methods can be used for finding the polymerase chain reaction in finding the viral medicine. The studies can be developed on collection of datasets that can provide a balance between public health and data privacy with AI interactions. The privacy of the data using blockchain technology can enable secured transactions of healthcare data and embedding AI for data analytics to predict the future of infectious spread.

Data Availability

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

Conflicts of Interest

None of the authors have any conflicts of interest.

References

- [1] B. Qian, X. Wang, N. Cao, H. Li, and Y. G. Jiang, "A relative similarity based method for interactive patient risk prediction," *Data Mining and Knowledge Discovery*, vol. 29, no. 4, pp. 1070–1093, 2015.
- [2] J. Ranjan and C. Foropon, "Big data analytics in building the competitive intelligence of organizations," *International Journal of Information Management*, vol. 56, p. 102231, 2021.
- [3] G. Ryu, A. Nasridinov, H. Rah, and K. H. Yoo, "Forecasts of the amount purchase pork meat by using structured and unstructured big data," *Agriculture*, vol. 10, no. 1, p. 21, 2020.
- [4] Z. Xu, J. Chou, X. S. Zhang et al., "Identifying sub-phenotypes of acute kidney injury using structured and unstructured electronic health record data with memory networks," *Journal of Biomedical Informatics*, vol. 102, p. 103361, 2020.
- [5] S. Singh and G. Jagdev, "Execution of structured and unstructured mining in automotive industry using Hortonworks sandbox," *SN Computer Science*, vol. 2, no. 4, pp. 1–25, 2021.
- [6] N. S. Spillane, M. R. Schick, K. T. Kirk-Provencher, D. C. Hill, J. Wyatt, and K. M. Jackson, "Structured and unstructured activities and alcohol and marijuana use in middle school: the role of availability and engagement," *Substance Use & Misuse*, vol. 55, no. 11, pp. 1765–1773, 2020.
- [7] Q. Zhu, X. Rui, Y. Li, Y. You, X. L. Sheng, and B. Lei, "Identification of four novel variants and determination of genotype-phenotype correlations for ABCA4 variants associated with inherited retinal degenerations," *Frontiers in cell and developmental biology*, vol. 9, 2021.
- [8] S. I. Ayon, M. M. Islam, and M. R. Hossain, "Coronary artery heart disease prediction: a comparative study of computational intelligence techniques," *IETE Journal of Research*, vol. 66, pp. 1–20, 2020.
- [9] Z. Song and S. Ren, "Product market competition and R&D investment: evidence from textual analysis on annual report of China's listed firms," *Asian Economics Letters*, vol. 1, no. 4, p. 17663, 2020.
- [10] J. V. Freitas-Jesus, L. Rodrigues, and F. G. Surita, "The experience of women infected by the COVID-19 during pregnancy in Brazil: a qualitative study protocol," *Reproductive Health*, vol. 17, no. 1, pp. 1–7, 2020.

- [11] S. N. Kasthurirathne, Y. A. Ho, and B. E. Dixon, "Public Health Analytics and Big Data," in *Public Health Informatics and Information Systems*, pp. 203–219, Springer, Cham, 2020.
- [12] A. Y. Odisho, B. Park, N. Altieri et al., "Natural language processing systems for pathology parsing in limited data environments with uncertainty estimation," *JAMIA open*, vol. 3, no. 3, pp. 431–438, 2020.
- [13] M. Petic, A. Gorea, and I. Ciobanu, "Important aspects in assessing the credibility of unstructured information," in *International Conference on Intelligent Information Systems*, pp. 151–156, 2020.
- [14] S. Talebian, G. G. Wallace, A. Schroeder, F. Stellacci, and J. Conde, "Nanotechnology-based disinfectants and sensors for SARS-CoV-2," *Nature Nanotechnology*, vol. 15, no. 8, pp. 618–621, 2020.
- [15] C. Weiss, M. Carriere, L. Fusco et al., "Toward nanotechnology-enabled approaches against the COVID-19 pandemic," *ACS Nano*, vol. 14, no. 6, pp. 6383–6406, 2020.
- [16] M. Vaculovicova, P. Michalek, S. Krizkova, M. Macka, and V. Adam, "Nanotechnology-based analytical approaches for detection of viruses," *Analytical Methods*, vol. 9, no. 16, pp. 2375–2391, 2017.