Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 2170839, 11 pages
https://doi.org/10.1155/2022/2170839

Research Article

LSGDM with Biogeography-Based Optimization (BBO) Model for Healthcare Applications

A. Harshavardhan,[1] Prasanthi Boyapati,[2] S. Neelakandan,[3] Alhassan Alolo Abdul-Rasheed Akeji,[4] Aditya Kumar Singh Pundir,[5] and Ranjan Walia[6]

[1]*Department of CSE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India*
[2]*Department of CSE, R.V.R & J.C College of Engineering, Guntur, India*
[3]*Department of CSE, R.M.K Engineering College, Chennai, India*
[4]*Department of Marketing and Corporate Strategy, Tamale Technical University, Tamale, Ghana*
[5]*Department of ECE, Arya College of Engineering and Information Technology, Jaipur, India*
[6]*Department of Electrical Engineering, Model Institute of Engineering and Technology, Jammu, India*

Correspondence should be addressed to Alhassan Alolo Abdul-Rasheed Akeji; aaakeji@tatu.edu.gh

Received 19 February 2022; Revised 26 March 2022; Accepted 7 April 2022; Published 30 April 2022

Academic Editor: Mohsen Ahmadi

Several studies aimed at improving healthcare management have shown that the importance of healthcare has grown in recent years. In the healthcare industry, effective decision-making requires multicriteria group decision-making. Simultaneously, big data analytics could be used to help with disease detection and healthcare delivery. Only a few previous studies on large-scale group decision-making (LSDGM) in the big data-driven healthcare Industry 4.0 have focused on this topic. The goal of this work is to improve healthcare management decision-making by developing a new MapReduce-based LSDGM model (MR-LSDGM) for the healthcare Industry 4.0 context. Clustering decision-makers (DM), modelling DM preferences, and classification are the three stages of the MR-LSDGM technique. Furthermore, the DMs are subdivided using a novel biogeography-based optimization (BBO) technique combined with fuzzy C-means (FCM). The subgroup preferences are then modelled using the two-tuple fuzzy linguistic representation (2TFLR) technique. The final classification method also includes a feature extractor based on long short-term memory (LSTM) and a classifier based on an ideal extreme learning machine (ELM). MapReduce is a data management platform used to handle massive amounts of data. A thorough set of experimental analyses is carried out, and the results are analysed using a variety of metrics.

1. Introduction

Recent technologies, such as big data, the internet of things (IoT), wearables, and so on, have a significant impact on society, healthcare organisations, and our daily lives. Big data plays an important role in obtaining the necessary data during the decision-making process. Big data is defined as a complex and massive volume of data derived from various sources and clinical data sets that provide critical information for patient treatment [1]. Furthermore, big data has the potential to improve healthcare operations through data-driven decision-making in the ambiguous environment of Industry 4.0. Big data analytics provide significant benefits for evaluating and assimilation of massive amounts of complex healthcare data. The medical system keeps track of the world's most pressing social and economic issues in order to find innovative solutions through technology and science. The Industry 4.0 model was first proposed in 2011, and it was initially referred to as the production or manufacturing process. While incorporating, medical services and Industry 4.0 are complementary methodologies. Furthermore, with the rise of big data and the

widespread use of electronic healthcare records of patients in healthcare organisations, chasing solutions to population medical problems is no longer viable. Using big data for better decision-making, on the other hand, poses some healthcare challenges.

In recent years, group decision-making has received a lot of attention in various areas of healthcare organisations [2]. The more severe the challenge, the more complex it can withstand the loss caused by a decision-making error. As a result, many civil organisations and government departments, as well as managers and experts in various fields, would be involved in decision-making. Based on the preferences of decision-makers, LSGDM selects sufficient alternates from a group of possible alternates [3]. When the number of decision-makers (DM) increases, the standard group decision-making problem transforms into the LSGDM problem. LSGDM problem-solving methods typically consist of four phases: (i) the cluster standardised individual decision matrices, (ii) standardising original individual decision metrics, (iii) selecting the best alternatives, and (iv) aggregating the cluster decision metrics.

Based on the conventional decision-making method, the current study significantly innovates by incorporating four factors: distinct decision-makers' preference data expression, attribute weight determination method, large-scale group clustering method, and large group preference data aggregation method [4]. One of the most common fields of study is large-scale group preference data aggregation. Despite the fact that the number of studies deliberating big data is steadily increasing, applications of large-scale group decision-making procedures in the context of big data studies and medical Industry 4.0 remain rare. This work creates a new MapReduce-based LSDGM model (MR-LSDGM) for the Industry 4.0 environment to improve decision-making in healthcare management. With the rapid advancement of information technology, as exemplified by the Internet, decision support systems will evolve toward socialisation in the era of big data. This is because DMs from various areas can be invited to collaborate on difficult issues on a single network platform. Simultaneously, we can conduct online voting on a particular item and perform automated statistical analysis. Additionally, it may analyse and research multitemporal and group events, such as those on e-commerce websites and search engines, to provide critical data support for qualitative decision-making. It appears to be worthwhile to develop large-scale group support tools to aid in decision-making, given that the big data era may contain a variety of data sources, including social media, mobile devices, and websites. Several researchers have developed software implementations of LSGDM, including the WTALGDM for LSGDM on energy network dispatch optimization, MENTOR for visualizing opinion evolution, and a multiagent system model for assisting with CRP. Other application fields, such as healthcare and engineering, require the use of these tools. A large number of simulations are run to demonstrate the improved results of the MR-LSDGM technique, and the experimental findings are examined using several metrics.

2. Related Works

Li and Wei [5] created an LSGDM model for making medical management decisions. For describing the decision data, the HFLTS is used. For clustering the DM into many subgroups, a clustering technique based on the ideal point is presented. The DM preference is then combined with the PDEHFLTS model to retain the decision data. A subgroup weight method is proposed for calculating the ranking weight based on the subgroup size and the presented hesitant entropy of PDEHFLTS. For large-scale GDM problems, Li et al. [6] used a fuzzy cluster analysis to integrate heterogeneous data. Fuzzy cluster analysis is used to divide large groups into smaller ones, and F-statistics are used to calculate the number of clusters required. The original data is kept depending on the degree of similarity. A consensus-building process is then used among these smaller groups to reach a common understanding. While other groups could not agree, a feedback system was devised to update the smaller GDM matrix, and the TOPSIS model was used to select the best option.

Song and Yuan [7] proposed a new GDM method based on arithmetic programming and employing IMGFLPR. As a result, a consensus procedure based on IMGFLPR is developed, while dynamic adaptation of expert weights is considered. Finally, problems with emergency plan election are solved using the presented method, which demonstrates the effective outcome of GDM. Wan et al. [8], inspired by multiplayer game concepts, proposed a two-step optimization algorithm that first maximises individual fulfilment while minimizing group conflict. The provided approach effectively saves decision-making time when it comes to ensuring the quality of LSGDM.

Hsu et al. [9] identified eight potential developments for providing a proper approach to the medical industry. The modified Z-DEMATEL method is used to build the mutually important relationship and prioritises this trend. By optimising the classic fuzzy number and representing the assessment environments' confidence under uncertainty, the Z-number technique improves the consistency of expert evaluation. Liu et al. [10] developed an LGDA approach for managing dependency in HRA based on the interval two-tuple linguistic variable and cluster analysis model. In addition, an expanded Muirhead mean operator was developed to determine the amounts of reliance between the activities of consecutive operators. Finally, empirical medical dependency analysis is used to demonstrate the applicability and effectiveness of the presented LGDA model [11].

Du and Shan [12] proposed a dynamic intelligent integration suggestion approach for product ideas. They began by creating one-of-a-kind product concept assessment condition systems that included both output and input criteria. The following section describes a step for static data combination and data extraction. Later, the fundamental likelihood assignments function is used as a data extraction approach to accurately reflect and effectively capture the

validity of experts' evaluations. Dursun et al. [13] proposed a fuzzy multicriteria GDM architecture based on the fuzzy integral and measure principle to evaluate the HCW treatment alternatives for Istanbul. In the case of the GDM problem, an expert consensus is required for the calculation method to be carried out correctly. The OWA operators are used in this work to aggregate DM opinions.

Pan et al. [14] concentrated on using dynamic programming to solve the large-scale GDM problem, in which the data is in the form of linguistic variables. Because the linguistic variable cannot be directly computed, the interval type-2 fuzzy set is used for encoding. The distinct similarity and distance models are then constructed concurrently in order to determine the relationship between the interval type-2 fuzzy set. Later, a dynamic programming approach based on clustering models was presented for clustering the DM from an overall perspective. Gao et al. [15] created a novel paradigm for selecting an appropriate physician in the index system that balances the 2D calculation result. The researchers created questionaries and conducted field research to bring the given technique closer to the actual situation in China. They then calculated the best outcome for the best medical services provided by doctors [16].

3. The Proposed MR-LSDGM Technique

The MapReduce tool is used in this study to create a new MR-LSDGM approach for the healthcare industry. The MR-LSDGM approach includes BBO-FCM-based DM clustering, 2TFLR-based preference modelling, and LSTM-OELM-based classification procedures. The proposed MR-LSDGM model is depicted in its entirety in Figure 1. The MapReduce tool is used by the MR-LSDGM approach to managing massive amounts of data in the healthcare industry. The sections that follow provide a more detailed explanation of these processes [17].

3.1. MapReduce. The primary goal of the Map procedure is to compute the geometric distance between cluster centres and sampling point data. Read the data from Hadoop Distributed File Systems (HDFS) and use the stated (value, key) pair input formats as Map function input values, where "key" denotes the sampling point data ID numbers and "value" means the entire data sampling point and then read the maximal consumption. The minimal distance approach would evaluate the major cluster centre, compute Euclidean distances between other cluster centres using sample point data, and integrate the membership degree (MD) [18].

The primary goal of Reduce functions, on the other hand, is to obtain a large number of Map function outputs. To begin, obtain the key values pair from the Map functions, where "key" represents the cluster centres and "value" represents the sampling point data equivalent to the cluster centres. The data sample from a number of distinct cluster centres is then merged, and a new cluster centre is evaluated. Finally, it is determined whether the geometric distance between the novel and equivalent cluster centres exceeds a predetermined threshold or whether the number of iterations exceeds that threshold.

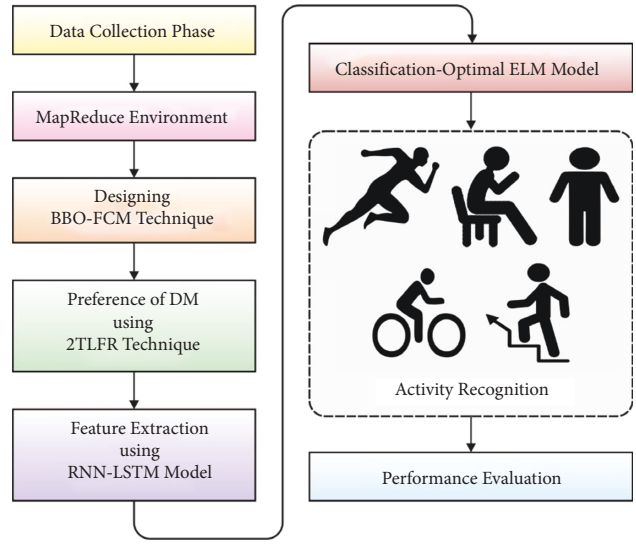

Figure 1: Overall process of MR-LSDGM model.

Despite outperforming traditional hard clustering algorithms in terms of clustering effects, fuzzy clustering algorithms have a few drawbacks. The current clustering algorithms are extremely sensitive to early clustering centres. Because the algorithms use the concept of gradual iteration, the objective functions are continuously reduced during the iteration. As a result, when the c clustering centre is arbitrarily chosen in each sample data set at first, the geometric distance that would produce the last clustering results for falling to the current optimum solution is smaller. To avoid situations in which the geometric distance between the arbitrarily chosen cluster centre is smaller, the minimum and maximum distance methods were used to determine the early cluster centre in this study.

3.2. Design of BBO-FCM Technique. In the beginning, the BBO-FCM approach is used to divide the DMs into subgroups. Every feature vector with a coefficient between [0, 1] belongs to one of the FCM clusters. Finally, the algorithm labels all of the data points (feature vectors) based on the maximum coefficient of these data points across all clusters. By minimizing the following equation, the cluster centre and fuzzy membership matrix are calculated.

$$\sum_{g=1}^{c} u_{i,j} = 1, \tag{1}$$

where u represents the sum of data; c indicates the quantity of clusters; $u_{g,j}$ signifies the fuzzy association of jth point to ith clusters; $d_{g,j}$ means the cluster centres and data point $l \in (1, \infty)$, a fuzzy weight factor that determines the quantity of fuzziness produced as a result In most cases, and $l = 2$ is chosen (it can be stated that this value of m does not generate optimum solutions for each problem).

Because of the constraints in (1), all points must completely allocate their memberships to each cluster [19]. The

fuzzy weight centre of gravity of the data is used to define the cluster centre (centroid)X.

$$v_j = n\sum_{j=1}^{n} u_{i,j}, x^l x_j. \tag{2}$$

As $u_{g,j}$ influences the calculation of the cluster centres v_i, data with more memberships would have a greater influence on the prototype position than data with fewer memberships. Because $u(g, j)$ has an effect on the calculation of the cluster centres v_i, it is necessary to consider it. The distance $d\_$ is used in the fuzzy C-means technique (g, j). Clustering using fuzzy logic (sometimes referred to as soft clustering or soft k-means) allows each individual data point to be assigned to more than one cluster. For fuzzy C-means approach, distance $d_{g,j}$ is determined by

$$d_{g,j}u^2 = \left\| x_j - v_i \right\|^2. \tag{3}$$

The cluster centre v_i represents the common value of that cluster, where the $u_{g,j}$ components of the association matrix denote the range where the data point x_j is related to its model. The minimalization of divide function (1) would derive the following equation:

$$\frac{d_{g,j}}{d_{g^j}}, x^{1/(l-1)} + u_{g,j} = \frac{1}{2}. \tag{4}$$

Equation (4) is defined in an iterative manner as the distance $d_{g,j}$ is based on membership $u_{i,k}$. The process to compute the FCM is given below:

(i) Opt for the number of cluster c, $2 \le c < n$; select m, $1 \le l < \infty$.Initialize$U^{(0)}$.

(ii) Compute the cluster centre v_i by (2).

(iii) Compute the novel partition matrix $U^{(1)}$ by (4).

(iv) Relate $U^{(j)}$ & $U^{(j+1)}$. When the variations of the MD $u_{k,i}$ computed by proper standards are smaller when compared to the provided threshold, end the process and return to step (2).

The BBO algorithm is used to define the optimal initial cluster centre of the FCM technique. The BBO has a population-based optimised technique that simulates the development and the balance of predator and prey in different ecosystems. According to research, the BBO produces better results than the other population-based techniques [20]. This technique utilises the BBO algorithm to select the optimal initial cluster centre to use for its initial cluster centre determination process. For each setting, the BBO uses an optimal population-based technique to simulate development and the balance between predators and prey. It has been discovered through research that the BBO generates superior results when compared to other population-based approaches. From one iteration to the next, a collection of solutions is retained, and all habitats send and receive inhabitants. The various habitats are determined by their immigration and emigration rates, which are probabilistically modified. An arbitrary number of habitats are occasionally mutated during all iterations. All of the solution parameters are now referred to as suitability index variables (SIV). Simon was the first to propose the concept of a biogeography-based optimization method. Using the scientific understanding of migration and the dispersion of species from one habitat to another, this method has been devised and tested. Each location has a habitat suitability index (HSI), which is based on the concepts of this algorithm and acts in a similar way to the fitness function in other population-centred algorithms. In addition, the suitability index variables refer to the independent factors that are used to determine the suitability index of a settlement (SIV).

The mathematical method of immigration (λ_k) and emigration (μ_k) is expressed as follows:

$$\lambda_k = I\left(1 - \frac{S_k}{S_{\max}}\right),$$
$$\mu_k = E\left(\frac{S_k}{S_{\max}}\right), \tag{5}$$

where I refers to the maximal rate of immigration, E defines the maximal rate of emigration, $S_{\max}$ implies the maximal number of habitats, and S_k represents the habitant count of k.

The following are the modifications to all habitats that improve the evaluation of BBO:

$$m(s) = m_{\max} \times \left(1 - \frac{P_n}{P_{\max}}\right), \tag{6}$$

where $m_{\max}$ represents the higher value of mutation determined as a user, $P_{\max}$ demonstrates the superior mutation probabilities of every habitat. and P_n refers to the mutation probabilities of n^{th} habitat that is given by

$$\acute{P}_n = \begin{cases} -(\lambda_n + \mu_n)P_n + \mu_{n+1}P_{n+1}, & n = 0; \\ -(\lambda_n + \mu_n)P_n + \mu_{n+1}P_{n+1} + \lambda_{n-1}P_{n-1}, & 1 \le n \le S_{\max} - 1 = 0; \\ -(\lambda_n + \mu_n)P_n + \lambda_{n-1}P_{n-1}, & n = S_{\max}. \end{cases} \tag{7}$$

At this point, $\underset{I:}{\varnothing} \longrightarrow \{H^n, HSI^n\}$ sets an ecosystem of habitat and calculates all equivalent HSIs, and $\Gamma = (n, m, \lambda, \tau, \Omega, M)$ describes the function that switches from one optimised cycle to the next. The six tuples of elements are described, where n implies the number of habitats, m refers to the number of SIVs, λ represents the rate of

immigration, τ demonstrates the rate of emigration, Ω refers to the migration function, and M indicates the mutation operator.

3.3. Modelling Preferences of DMs Using 2TLFR Technique. Once the DMs have been clustered, their perspectives can be defined and fused using the 2TLFR technique to retain as much decision information as possible. Decision-making in healthcare is based on dynamic conditions and ambiguous information, and most decision-makers prefer linguistic variables or fuzzy values over hard numbers. In the two-tuple linguistic depiction method, the data measured in the linguistic hierarchy term set could be unified with no data loss.

Definition 1. $S = \{s_0, s_1, \ldots, s_g\}$ represents a linguistic term set; $\beta\epsilon[0, g]$ denotes the outcome of an aggregation index of a group of labels measured in S. (s_γ, α) indicates linguistic two tuples; $s_\gamma \epsilon S$ and $\alpha\epsilon[-0.5, 0.5]$; s_γ characterises the linguistic label of the data; and α means the mathematical value that expresses the values of the translation from the original results β to the nearest index label γ in S, namely, the symbolic translation.

The following function converts mathematical numbers and linguistic two tuples. They can convert mathematical values to linguistic two tuples using (9) [8].

$$\Delta\colon [0, g] \longrightarrow s\times, \tag{8}$$

$$\Delta(\beta) = \begin{cases} s_\gamma, & \gamma = \text{round}(\beta), \\ \alpha = \beta - r, & \alpha \in [-0.5, 0.5]. \end{cases} \tag{9}$$

Using the eq., they can convert a linguistic phrase to a real value (between 0 and g).

$$\begin{aligned} \Delta^{-1}\colon s \times [-0.5, 0.5] &\longrightarrow [0, g] \\ \Delta^{-1}(s_\gamma, \alpha) &= \gamma + \alpha = \beta \end{aligned}. \tag{10}$$

To further unify the dimension, they could use the eq. to map the linguistic term between zero and one.

$$\begin{aligned} &\Delta^{-1}\colon s \times [-0.5, 0.5] \longrightarrow [0, 1], \\ &\Delta^{-1}(s_\gamma, \alpha) = \frac{\gamma + \alpha}{g}. \end{aligned} \tag{11}$$

3.4. Automated Disease Classification Model. Finally, the disease classification process is divided into three stages: feature extraction using LSTM, classification using ELM, and parameter tuning using tree growth algorithm (TGA). As previously stated, convolution models can work on a single image and transform it from input pixel to matrix/vector representation. Current CNN pretrained models are used for feature extraction. The main goal is that CNN may not be trained, but training may be provided by the BP errors from the LSTM-DL classifiers via CNN multiple input images. Convolutional neural networks (CNNs) are used because of their improved transferability. Knowledge of this cutting-edge technology will benefit not just researchers who use CNN for radiology and medical imaging jobs but also clinical radiologists, since deep learning may influence their practise in the near future. Following CNN training, medical professionals or computer-aided detection (CADe) systems can specify the target lesions in medical pictures during the deployment phase. Figure 2 depicts a general LSTM cell. The LSTM cell contains various gates and parameters that control the behaviour of each memory cell. Every cell state is governed by the activation function of gates. For different types of gates, the input value is fed into the input gate (I), forget gate (f), activation vector (c), and output gate (o).

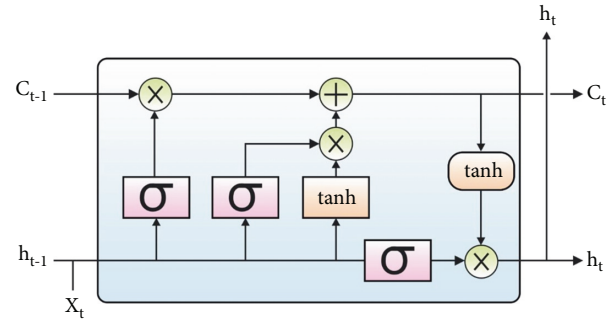

Figure 2: LSTM structure.

$$\begin{aligned} j_t &= \epsilon\, j\left(w_{pj}p_t + w_{hj}h_{t-1} + w_{aj}a_{t-1} + bj\right), \\ f_t &= \epsilon\, f\left(wPp_t + w_h h_{t-1} + w_{af}a_{-1} + bf\right), \\ a_t &= f_t a_{t-1} + j_t \epsilon_a\left(w_{pc}p_t + w_{ha}h_{t-1} + ba\right), \\ 0_t &= \epsilon_o\left(w_{po}p_t + w_{hot-1}h + w_{ao}a_{t-1} + bo\right), \\ h_t &= 0_t \in h(a_t), \end{aligned} \tag{12}$$

where $w_{pj}, w_{hj}, w_{aj}, w_{hf}, w_{af}, w_{pa}, w_{ha}, w_{po}$, and w_{ao} denote weight (input weight, hidden weight, output weight, etc.) and b_j, b_f, b_a, and b_0 indicate the bias weight [21].

Each time step includes a single CNN series and an LSTM model. As a single step, CNN could be passed and used on every output to the LSTM input image for all input images. The result could be achieved by folding up a CNN input framework with multiple layers in a time distribution method. The same layer is used multiple times to achieve a similar result. To determine the presence of diseases, the extracted features are fed into the ELM classifier. Assume a training data $A\ \{(x_i, t_i)\}_{i=1}^N$, the output functions of SLFN using L hidden neuron could be determined by

$$f(x_i) = \sum_{j=1}^{L} \beta_j h_j\left(a_j, b_j, x_i\right) = h(x_i)\beta, i = 1, \ldots, N, \tag{13}$$

where $\beta = \beta_1, \ldots, \beta_L^T$ denotes the output weight matrix and $h(x_i) = [h_1(a_1, b_1, x_i), \ldots, h_j(a_L, b_L, x_i)]$ represents the network output equivalent to the training samples x_i. $h_j(\cdot)$ indicates a nonlinear piecewise continuous function and $a_j \in R^d$, and $b_j \in R\,(j = 1, 2, \ldots, L)$ means a parameter of jth hidden node. The training network is to discover appropriate network parameters for minimizing the error functions $\|H\beta - T\|_2$, where

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix}, \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}. \tag{14}$$

It represents an SLFN using an L hidden neuron and denotes the hidden, output matrix, and target output.

ELM uses arbitrary hidden node parameters and the tune-free trained approach to FFNN instead of iteratively upgrading network parameters as in traditional gradient descent algorithms. ELM is flexible because it employs a hidden activation function, as demonstrated by the universal approximate ability theorem. Almost any nonlinear piecewise continuous function and its linear combination perform well in the ELM algorithm [22]. The extreme learning machine (ELM) is a fast convergent training method for single hidden layer feedforward neural networks (SLFNs). This type of SLFN allows for faster convergence training and avoids the need for many iterations to update the hidden layer weights. Compared to other classical learning algorithms in applications with increasing noise, ELM appears to outperform ELM in regression and classification tests. With a single hidden layer of neurons and random feature mapping, an ELM model learns quicker than other models. High dimensions and large data sets have aroused substantial scholarly interest in the low computing complexity.

The TGA is used to optimise the ELM model's parameter computation, resulting in improved overall classification performance. The TGA approach is stimulated by the competition between trees in the forest. A tree's attention is divided between food and sunlight. Exploration and exploitation are the two major stages of the approach. During the exploration stage, the tree moves toward the sunlight, allowing it to investigate new locations. The tree is now fulfilled with light in the exploitation stage, and thus, it moves towards better nutrients in the root, as it moves towards the global/local optimal. The forest's tree population is classified into four types.

The first group of trees has found a light source, and they can now compete for food. To compete with light, the tree in the second group switches to the two optimal options that are closest to it. In the third group, a new tree is planted in place of the worst tree. Finally, an optimal tree is used to create a novel plant [23]. Initially, this approach arbitrarily creates the early population of the tree (solution) within the lower and upper bounds, where the fitness values for all solutions are calculated. The following is how the early population is produced.

$$x_{i,j} = \min_j + \text{rand} \cdot \left(\max_j - \min_j\right), \tag{15}$$

where $x_{i,j}$ is the jth variable of ith solution of the population, rand represents an arbitrary value derived by the uniform distribution, and $\min_j$ and $\max_j$ indicates lower bound and upper bound of jth variable, respectively.

Next, the population is arranged based on the fitness values, and the present optimal solution at the jth iterations is established. The global optimal solutions are represented as T_{GB}^j. The optimal solution is allocated to the initial group (N_1), and the solution from this population carries out the local search as follows:

$$T_i^{j+1} = \frac{T_i^j}{\theta} + r \times T_i^j, \tag{16}$$

where T_i^{j+1} represents the novel ith solution and T_i^j is the ith solution in jth iteration. θ represents the rate of power decreases, and r specifies an arbitrary number between [0,1].

When the novel solution has a higher fitness value than the current solution, greedy selections are used to find it. The novel solution either replaces the current one or keeps the current solutions for the next generation.

The second optimal solution is allocated to N_2 subpopulation. All solutions from the N_2 group must be shifted to the two nearest solutions (from the initial and second subpopulation) at distinct α angles. The Euclidean distance is used to measure the distance between two solutions.

$$d_i = \left(\sum_{i=1}^{N_1+N_2} \left(T_{N_2}^j - T_i^j\right)^2\right)^{1/2}, \quad d_i = \begin{cases} d_i \text{ if} T_{N_2}^j \neq T_i^j, \\ \infty \text{ if} T_{N_2}^j = T_i^j, \end{cases} \tag{17}$$

where d_j denotes the distance of ith solution, The trees that exist now are depicted as $T_{N_2}^j$, and the ith solution in the population are signified as T_i^j.

The poorest solution from the population is found in the third subpopulation, N_3. This solution is calculated by replacing it with a recent arbitrary solution.

$$N_3 = N - N_1 - N_2, \tag{18}$$

where the population sizes are represented by $N.N_1$ and N_2 are the first and second subpopulation, respectively. Next, the novel population (N) is determined by adding the initial groups N_1, N_2, and N_3.

$$N = N_1 + N_2 + N_3. \tag{19}$$

The last group N_4 includes arbitrary novel results.N_4 is the final group of entire set outcomes which contains arbitrary novel findings.. Using mask operators, the population adapts an optimal solution from the initial group (N), and the adapted solution is fused. The fitness values are used to organise the novel population, and the best N solution is chosen for the next iteration. The procedure is repeated until the desired result is obtained. Finally, the best solution is determined.

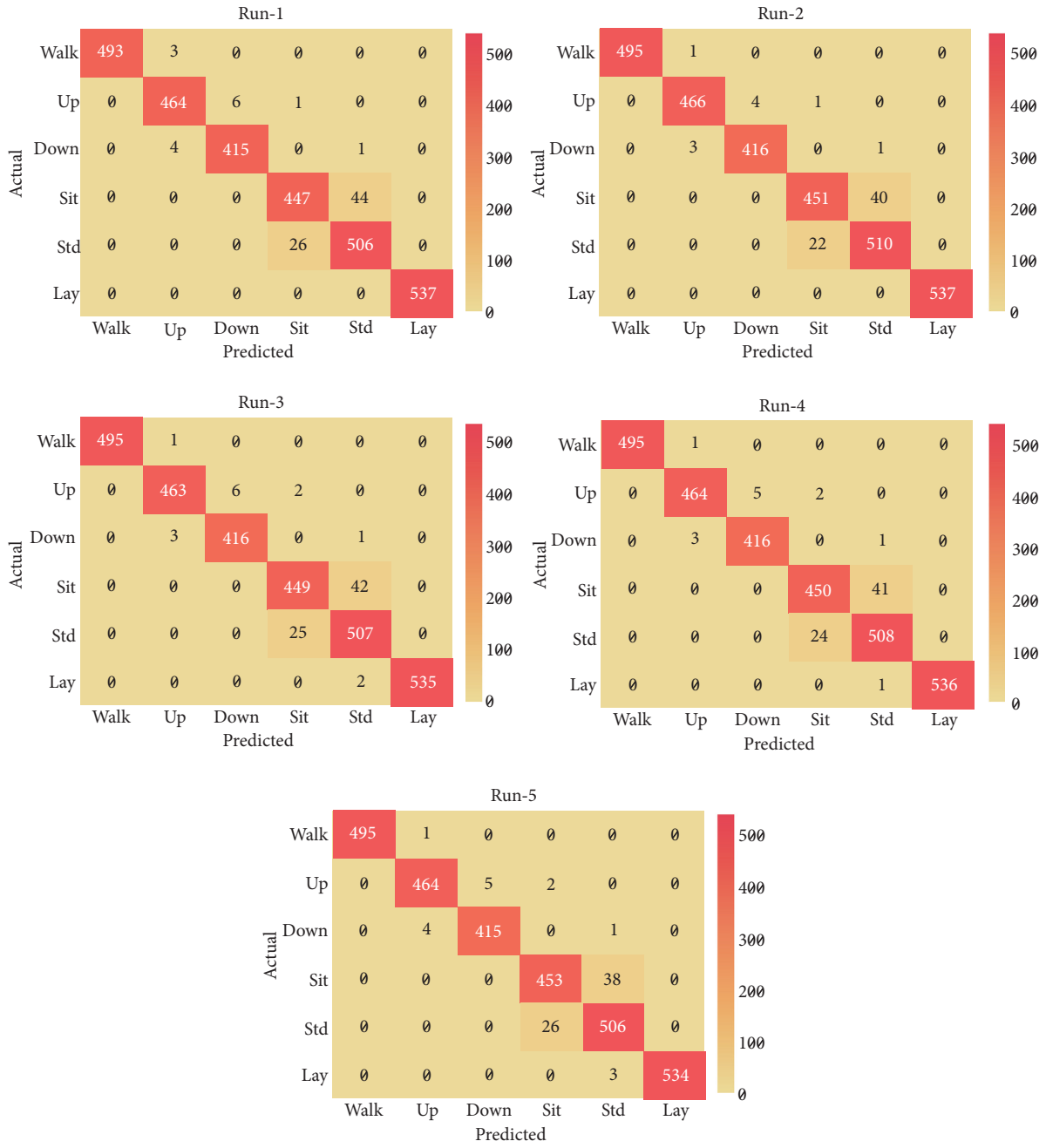

Figure 3: Confusion matrix analysis of MR-LSDGM model.

4. Performance Validation

The performance of the MR-LSDGM approach is investigated in this section using the benchmark activity recognition data set from the UCI repository [24]. The data set contains information on 30 people, each with 561 attributes. The data set contains 496 instances from the Walk class, 471 instances from the Up class, 420 instances from the Down class, 491 instances from the Sitting class, 532 instances from the Standing class, and 537 instances from the Lying class.

After five repetitions, the MR-LSDGM approach produced a collection of five confusion matrices, as shown in Figure 3. The graph shows that the MR-LSDGM method yielded the best possible result in each execution run [25]. For example, the MR-LSDGM technique classified 493 instances as Walk, 464 instances as Up, 415 instances as Down, 447 instances as Sit, 506 instances as Stand, and 537 instances as Lay under run-1. Similarly, the MR-LSDGM approach classified 495 instances as Walk, 466 instances as Up, 416 instances as Down, 451 instances as Sit, 510 instances as Stand, and 537 instances as Lay in run-2. Similarly, the MR-LSDGM method classified 495 instances as Walk, 464 instances as Up, 416 instances as Down, 450 instances as Sit, 508 instances as Stand, and 536 instances as Lay under run-4. Furthermore, under run-5, the MR-LSDGM algorithm classified 495 instances as Walk, 464 instances as Up,

Table 1: Result analysis of MR-LSDGM technique under different runs.

No. of runs	Methods	Sensitivity	Specificity	Precision	Accuracy	*F*-score
Run-1	Walk	0.994	1.000	1.000	0.999	0.997
	Up	0.985	0.997	0.985	0.995	0.985
	Down	0.988	0.998	0.986	0.996	0.987
	Sit	0.910	0.989	0.943	0.976	0.926
	Std	0.951	0.981	0.918	0.976	0.934
	Lay	1.000	1.000	1.000	1.000	1.000
	Average	**0.971**	**0.994**	**0.972**	**0.990**	**0.972**
Run-2	Walk	0.998	1.000	1.000	1.000	0.999
	Up	0.989	0.998	0.992	0.997	0.990
	Down	0.991	0.998	0.991	0.997	0.991
	Sit	0.919	0.991	0.952	0.979	0.935
	Std	0.959	0.983	0.926	0.979	0.942
	Lay	1.000	1.000	1.000	1.000	1.000
	Average	**0.976**	**0.995**	**0.977**	**0.992**	**0.976**
Run-3	Walk	0.998	1.000	1.000	1.000	0.999
	Up	0.983	0.998	0.991	0.996	0.987
	Down	0.991	0.998	0.986	0.997	0.988
	Sit	0.915	0.989	0.943	0.977	0.929
	Std	0.953	0.981	0.919	0.976	0.935
	Lay	0.996	1.000	1.000	0.999	0.998
	Average	**0.973**	**0.994**	**0.973**	**0.991**	**0.973**
Run-4	Walk	0.998	1.000	1.000	1.000	0.999
	Up	0.985	0.998	0.992	0.996	0.988
	Down	0.991	0.998	0.988	0.997	0.989
	Sit	0.917	0.989	0.945	0.977	0.931
	Std	0.955	0.982	0.922	0.977	0.938
	Lay	0.998	1.000	1.000	1.000	0.999
	Average	**0.974**	**0.995**	**0.975**	**0.991**	**0.974**
Run-5	Walk	0.998	1.000	1.000	1.000	0.999
	Up	0.985	0.998	0.989	0.996	0.987
	Down	0.988	0.998	0.988	0.997	0.988
	Sit	0.923	0.989	0.942	0.978	0.932
	Std	0.951	0.983	0.923	0.977	0.937
	Lay	0.994	1.000	1.000	0.999	0.997
	Average	**0.973**	**0.995**	**0.974**	**0.991**	**0.973**

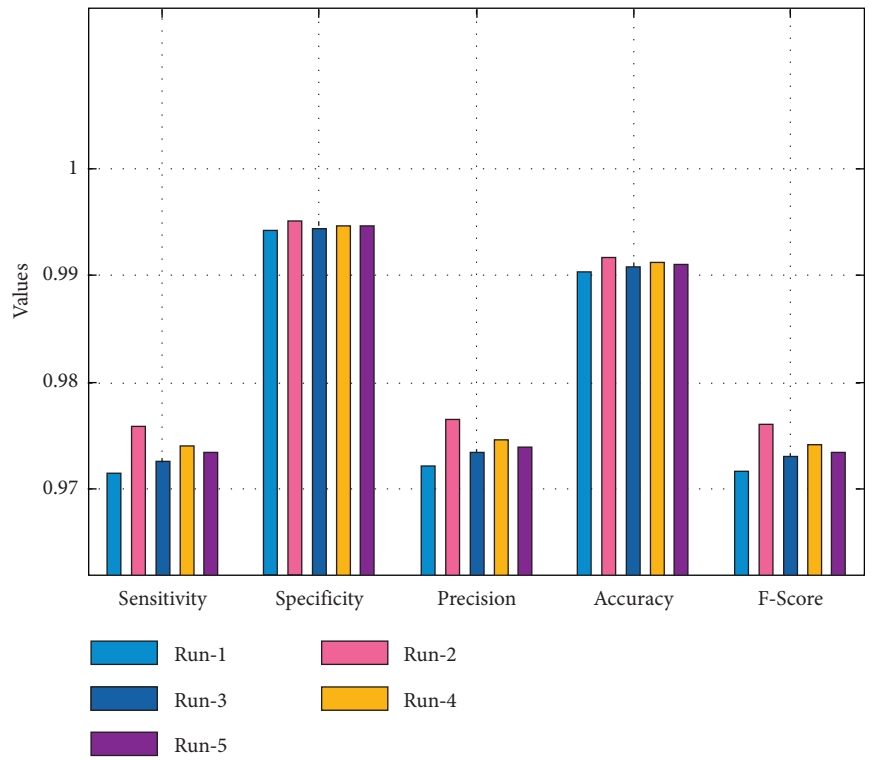

Figure 4: Result analysis of MR-LSDGM model with different measures.

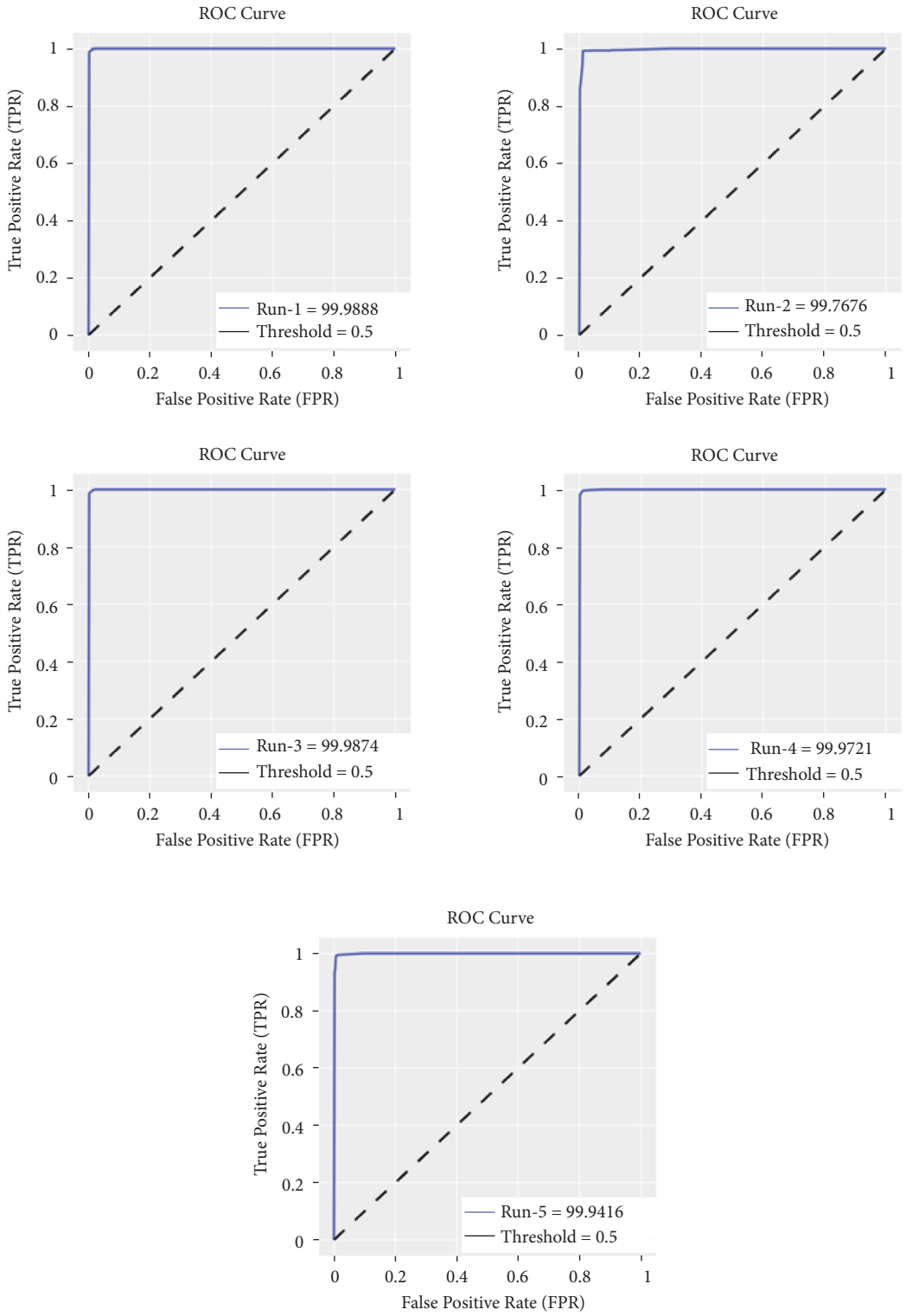

Figure 5: ROC analysis of MR-LSDGM model.

415 instances as Down, 453, Sit, 506, Stand, and 534 instances as Lay [26].

The classification result analysis of the MR-LSDGM technique under varying execution runs is reported in Table 1 and Figure 4. The MR-LSDGM technique has resulted in superior performance across all runs, as shown in Table 1. For example, the MR-LSDGM technique achieved maximum performance with run-1, with an average sensitivity of 0.971, specificity of 0.994, precision of 0.972, accuracy of 0.990, and *F*-score of 0.972. The MR-LSDGM method also performed optimally in run-2, with an average sensitivity of 0.976, specificity of 0.995, precision of 0.977, accuracy of

Table 2: Comparative accuracy analysis of MR-LSDGM with other techniques.

Methods	Accuracy	Precision	ROC
CNN-2016	0.9375	0.9554	0.9454
CNN-2018	0.9531	0.9638	0.9538
CNN-SF	0.9763	0.9655	0.9555
CNN-LSTM	0.9580	0.9755	0.9655
Lightweight CNN	0.9627	0.9822	0.9722
CNN-BiLSTM	0.9705	0.9852	0.9752
MR-LSDGM	0.9910	0.9925	0.9825

0.992, and *F*-score of 0.976. Furthermore, with run-3, the MR-LSDGM method achieved an average sensitivity of 0.973, specificity of 0.994, precision of 0.973, accuracy of 0.991, and *F*-score of 0.973. With run-5, the MR-LSDGM approach improved efficiency, achieving an average sensitivity of 0.973, specificity of 0.995, precision of 0.974, accuracy of 0.991, and *F*-score of 0.973.

Figure 5 depicts the ROC analysis of the MR-LSDGM method on the applied data set under various runs [27]. According to the results, the MR-LSDGM approach had the highest ROC value in every run. For example, in run-1, the MR-LSDGM technique achieved an increased ROC of 99.9888. In line with run-2, the MR-LSDGM method has a better ROC of 99.7676. The MR-LSDGM methodology then achieved a maximum ROC of 99.9874 in run-3. Concurrently, the MR-LSDGM technique achieved a superior ROC of 99.9721 in run-4. Finally, under run-5, the MR-LSDGM method achieved a maximum ROC of 99.9416 [28].

An extended comparison analysis is provided in Table 2 [25] to demonstrate the improved performance of the MR-LSDGM technique. With accuracy of 0.9375 and 0.9531, respectively, the CNN-2016 and CC-2018 approaches produced ineffective results [29]. At the same time, the CNN-LSTM and lightweight CNN approaches improved their accuracy to 0.9627 and 0.958, respectively. Furthermore, the CNN-BiLSTM and CNN-SF approaches have acceptable accuracy values of 0.9705 and 0.9763, respectively. In contrast, the proposed MR-LSDGM approach achieved an effective performance of 0.991 [30].

As evidenced by the tables and statistics above, the MR-LSDGM technique is clearly more effective than the other procedures.

4.1. Discussion. The healthcare IoT data sets and performance criteria for the proposed MR-LSGDM strategy are briefly outlined in this section [31]. The complete approach was developed using the MATLAB 2021a tool on a Core i3-3110M processor running Windows 8 with 2 GB RAM, and it was tested on 8 healthcare IoT data sets (Table 1) [32]. Over 30 separate runs, the new BBO-FCM approach was compared to existing algorithms such as CNN 2016, CNN 2018, CNN-SF, CNN-LSTM, lightweight CNN, and CNN-BiLSTM in terms of intracluster distance, purity index, standard deviation, root mean square error, accuracy, and *F*-measure [33].

5. Conclusion

The MapReduce tool is used in this study to create a new MR-LSDGM approach for the healthcare sector. The MR-LSDGM approach includes BBO-FCM-based DM clustering, 2TFLR-based preference modelling, and LSTM-OELM-based classification procedures. To manage big data in the healthcare sector, the MR-LSDGM technique employs the MapReduce tool. Furthermore, the design of the BBO algorithm for determining the primary cluster centres of the FCM technique, as well as parameter optimization of ELM using the TGA technique, contribute to improved overall classification results. A large number of simulations are run to demonstrate the improved outcomes of the MR-LSDGM technique, and the experimental results are examined using several metrics. According to the simulation results, the MR-LSDGM methodology outperformed the other methods. In the future, the model presented here could be used in telemedicine applications to help patients in remote areas.

Data Availability

The manuscript contains all of the data.

Conflicts of Interest

The authors declare that they do not have any conflicts of interest.

References

[1] S. Neelakandan and D. Paulraj, "A gradient boosted decision tree-based sentiment classification of twitter data," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 18, no. 4, Article ID 2050027, pp. 1-27, 2020.

[2] S. Neelakandan, S. Divyabharathi, S. Rahini, and G. Vijayalakshmi, "Large scale optimization to minimize network traffic using MapReduce in big data applications," in *Proceedings of the International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*, pp. 193–199, Melmaruvathur, India, April 2016.

[3] A. Labella, Y. Liu, R. M. Rodríguez, and L. Martínez, "Analyzing the performance of classical consensus models in large scale group decision making: a comparative study," *Applied Soft Computing*, vol. 67, pp. 677–690, 2018.

[4] T. Kavitha, P. P. Mathai, C. Karthikeyan et al., "Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 1, pp. 113–129, 2021.

[5] S. Li and C. Wei, "A large scale group decision making approach in healthcare service based on sub-group weighting model and hesitant fuzzy linguistic information," *Computers & Industrial Engineering*, vol. 144, Article ID 106444, 2020.

[6] G. Li, G. Kou, and Y. Peng, "Heterogeneous large-scale group decision making using fuzzy cluster analysis and its application to emergency response plan selection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 324, pp. 1–13, 2021.

[7] Y. Song and F. Yuan, "An interactive consensus model for

group decision-making with incomplete multi-granular 2-tuple fuzzy linguistic preference relations," *Mathematical Problems in Engineering*, vol. 2018, Article ID 8197414, 11 pages, 2018.

[8] Q. Wan, X. Xu, X. Chen, and J. Zhuang, "A two-stage optimization model for large-scale group decision-making in disaster management: minimizing group conflict and maximizing individual satisfaction," *Group Decision and Negotiation*, vol. 29, no. 5, pp. 901–921, 2020.

[9] W. C. J. Hsu, J. J. Liou, and H. W. Lo, "A group decision-making approach for exploring trends in the development of the healthcare industry in Taiwan," *Decision Support Systems*, vol. 141, Article ID 113447, 2021.

[10] H. C. Liu, Z. Li, J. Q. Zhang, and X. Y. You, "A large group decision making approach for dependence assessment in human reliability analysis," *Reliability Engineering & System Safety*, vol. 176, pp. 135–144, 2018.

[11] H. Singh, D. Ramya, R. Saravanakumar et al., "Artificial intelligence based quality of transmission predictive model for cognitive optical networks," *Optik*, vol. 257, Article ID 168789, 2022.

[12] Y. W. Du and Y. K. Shan, "A Dynamic Intelligent Recommendation Method Based on the Analytical ER Rule for Evaluating Product Ideas in Large-Scale Group Decision-Making," *Group Decision and Negotiation*, vol. 30, pp. 1–21, 2020.

[13] M. Dursun, E. E. Karsak, and M. A. Karadayi, "A fuzzy multi-criteria group decision making framework for evaluating health-care waste disposal alternatives," *Expert Systems with Applications*, vol. 38, no. 9, Article ID 11453, 2011.

[14] X. Pan, Y. Wang, S. He, and K. S. Chin, "A dynamic programming algorithm-based clustering model and its application to interval type-2 fuzzy large-scale group decision making problem," *IEEE Transactions on Fuzzy Systems*, vol. 30, 2020.

[15] Y. Gao, Y. Du, H. Liang, and B. Sun, "Large Group Decision-Making Approach Based on Stochastic MULTIMOORA: An Application of Doctor Evaluation in Healthcare Service," *Complexity*, vol. 2018, Article ID 5409405, 13 pages, 2018.

[16] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes–Ortiz, "A public domain dataset for human activity recognition using smartphones," in Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium, April 2013.

[17] O. Nafea, W. Abdul, G. Muhammad, and M. Alsulaiman, "Sensor-based human activity recognition with spatio-temporal deep learning," *Sensors*, vol. 21, no. 6, p. 2141, 2021.

[18] C. Zhen, "Using Big Data Fuzzy K-Means Clustering and Information Fusion Algorithm in English Teaching Ability Evaluation," *Complexity*, vol. 2021, Article ID 5554444, 9 pages, 2021.

[19] S. M. Zabihi and M. R. Akbarzadeh-T, "Generalized Fuzzy C-Means Clustering with Improved Fuzzy Partitions and Shadowed Sets," *International Scholarly Research Notices*, vol. 2012, Article ID 929085, 6 pages, 2012.

[20] X. Pu, S. Chen, X. Yu, and L. Zhang, "Developing a Novel Hybrid Biogeography-Based Optimization Algorithm for Multilayer Perceptron Training under Big Data challenge," *Scientific Programming*, vol. 2018, Article ID 2943290, 7 pages, 2018.

[21] S. Rani, H. Babbar, S. Coleman, A. Singh, and H. M. Aljahdali, "An efficient and Lightweight deep learning model for human activity recognition using smartphones," *Sensors*, vol. 21, no. 11, p. 3845, 2021.

[22] S. Neelakandan, S. Tripathi, V. B. Devi, I. Bhardwaj, and N. Arulkumar, "IoT-based traffic prediction and traffic signal control system for smart city," *Soft Computing*, vol. 25, no. 18, Article ID 12241, 2021.

[23] N. Bacanin, T. Bezdan, E. Tuba, I. Strumberger, and M. Tuba, "Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics," *Algorithms*, vol. 13, no. 3, p. 67, 2020.

[24] R. Kamalraj, S. Neelakandan, M. Ranjith Kumar, V. Chandra Shekhar Rao, R. Anand, and H. Singh, "Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm," *Measurement*, vol. 183, Article ID 109804, 2021.

[25] G. Reshma, C. Al-Atroshi, V. Kumar Nassa et al., "Deep learning-based skin lesion diagnosis model using dermoscopic images," *Intelligent Automation & Soft Computing*, vol. 31, no. 1, pp. 621–634, 2022.

[26] S. Neelakandan, A. Arun, R. Ram Bhukya, B. Hardas, T. Ch Anil Kumar, and M. Ashok, "An automated word embedding with parameter tuned model for web crawling," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1617–1632, 2022.

[27] S. Satpathy, M. Prakash, S. C. Sengupta Aditya, and K. D. Bhattacaryya Bidyut, "Design a FPGA, fuzzy based, insolent method for prediction of multi-diseases in rural area," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 7039–7046, 2019.

[28] S. Satpathy, P. Mohan, and S. Debbarma, "A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA," *The Journal of Supercomputing*, vol. 76, no. 8, pp. 5849–5861, 2020.

[29] M. Sundaram, "An analysis of air compressor fault diagnosis using machine learning technique," *Journal of Mechanics of Continua and Mathematical Sciences*, November - December, vol. 14, no. No.-6, pp. 13–27, 2019.

[30] J. Rene Beulah, L. Prathiba, G. L. N. Murthy, E. Fantin Irudaya Raj, and N. Arulkumar, "Blockchain with deep learning-enabled secure healthcare data transmission and diagnostic model," *International Journal of Modeling, Simulation, and Scientific Computing*, 2022.

[31] C. P. D. Cyril, J. R. Beulah, N. Subramani, and D. Sivabalaselvamani, "An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM," *Concurrent Engineering*, vol. 29, no. 4, pp. 386–395, 2021.

[32] D. Venu, A. V. R. Mayuri, S. Neelakandan, G. Murthy, N. Arulkumar, and N. Shelke, "An efficient low complexity compression based optimal homomorphic encryption for secure fiber optic communication," *Optik*, vol. 252, 2022.

[33] P. Asha, L. Natrayan, B. T. Geetha et al., "IoT enabled environmental toxicology for air pollution monitoring using AI techniques," *Environmental Research*, vol. 2022, Article ID 112574, 2022.