Early Detection and Classification of Biomedical Atherosclerosis using Feature Subset Selection with Optimal Deep Learning Model

Mohammed Abdalla[1]

[1]Faculty of Computers and Artificial Intelligence, Beni-Suef University, Egypt
[1]mohammed.a.youssif@fcis.bsu.edu.eg

Abstract

Atherosclerosis is a pathological disorder which grows progressively over the years and might result in cardiac arrest, stroke, or a peripheral vascular ailment, based on the region of existence in the human arterial network. Earlier detection and classification of atherosclerosis becomes crucial in mitigating the severity of the disease and death rate. The recently developed artificial intelligence (AI) methods like machine learning (ML) and deep learning (DL) models ensured their efficiency in the design of medical decision support systems (MDSS). In this view, this paper presents an early detection and classification of atherosclerosis using feature subset selection with optimal deep learning (EDCA-FSSODL) model. The proposed EDCA-FSSODL technique targets to decrease the dimensionality of the features and diagnose atherosclerosis. The proposed EDCA-FSSODL technique derives a novel binary chaotic flower pollination algorithm (BCFPA) based feature selection to eliminate the curse of dimensionality problem. In addition, Manta ray Foraging Optimization (MRFO) with bidirectional long short-term memory (BiLSTM) method is implemented for classifying and recognizing atherosclerosis disease. Moreover, the utilization of MRFO method assists in appropriate tuning of the hyperparameters of the BiLSTM technique and thus boosts the overall recognition performance. To portray the improved accomplishment of the EDCA-FSSODL technique, a sequence of simulations occurs by employing two standard datasets and validate the outputs by means of different measures. An overall relative research highlighted the supremacy of the EDCA-FSSODL technique over the other approaches.

Keywords

Atherosclerosis, Medical Decision Support System, Deep Learning, Metaheuristics, BiLSTM, Feature selection

1. Introduction

Atherosclerosis relates to the artery's blocking because of plaque settlement in its internal walls [1]. Cardiovascular disease (CVD) is the major source of demises around the globe, and 71% of this CVD occurs because of atherosclerosis,

according to international healthcare agency [2]. These pathological conditions can be prevented by tackling behavioural risk factors like smoking, lack of exercise, alcohol abuse, sedentary lifestyle, and so on. But in such cases, it requires medical interference, or else it could be demonstrated as fatal [3]. As well, the utmost unusual fact about CVD is that, it consumes more than a few years beforehand its symptom become predominant. The study is focused on detecting these pathological conditions in an earlier stage. There are different methods to accomplish the abovementioned objectives and are categorized into mathematical modeling, imaging, and signal processing. Indeed, such problems could be efficiently resolved when CVD is detected at an earlier stage, for starting an early interference. Invasive coronary angiography has long been the benchmark for the diagnosis of CVD [4]. This process depends on the support of percutaneous catheterization technique. It is an invasive process, which has the possibility of iatrogenic infection, vascular injury, etc. Hence, researcher begins to design efficient diagnostic and non-invasive tools for CVD. With the help of healthcare system, a massive number of healthcare information is gathered, involving laboratory examination data, medical history, and auxiliary examination [5]. These messages are utilized for establishing risk predictive systems which could assist in predicting and preventing upcoming CVD existence.

For clinical diagnoses of this disease, manual extracting of valuable data from patient archives is challenging. Therefore, the crucialness of constructing and establishing a Medical Diagnosis Support Scheme (MDSS) for the automation process of the affected individuals' classification and anticipation of CVD [6]. But clinical diagnosing research requires greater accuracy and effectiveness to make an enhanced medical conclusion. Though conventional MDSS has proved their capability for covering almost all the diagnosing complications, they grant a lesser rate of accuracy and cannot offer precise diagnosing [7]. In the previous years, diagnosing systems and clinical therapy using artificial intelligence (AI) and machine learning (ML) approaches have attained significant attention. Hence, the current research has influenced the field of research study like finances, applied sciences, medical, and biology applications. Therefore, several researches have been suggested to develop MDSS for anticipating or classifying persons with CVD for improving medical care [8]. Detecting CVD by ML algorithm is gaining considerable interest from the researcher [9]. Decision tree (DT), support vector machine (SVM), artificial neural network (ANN), and random forest (RF) are the more commonly adapted ML techniques for CVD study. It employs most of the datasets reported in the research with ease of use, low computational burden, and good performance.

This paper presents an early detection and classification of atherosclerosis using feature subset selection with optimal deep learning (EDCA-FSSODL) model. The proposed EDCA-FSSODL technique targets to decrease the dimensionality of the features and diagnose atherosclerosis. The proposed EDCA-FSSODL technique derives a novel binary chaotic flower pollination algorithm (BCFPA) based feature selection to eliminate the curse of dimensionality problem. In addition, Manta ray Foraging Optimization (MRFO) with bidirectional long short-term memory (BiLSTM) method is implemented for classifying and recognizing atherosclerosis disease. Moreover, the utilization of MRFO method assists in appropriate tuning of the hyperparameters of the BiLSTM technique and thus boosts the overall recognition performance. To portray the improved accomplishment of the EDCA-FSSODL technique, a sequence of simulations occurs by employing two standard datasets and validate the outputs by means of different measures.

2. Related Work

The current segment performs a detailed review of current MDSS tools for the detection and classification of atherosclerosis. Jain et al. [10] presented the concept of online model parameter tracking by exploiting an unscented Kalman filter (UKF) based architecture that will assist in monitoring the abovementioned pathological conditions. Upretee and Yüksel [11] presented a method-based classifying of a single time-changing spectral factor through DL approaches for the diagnoses of CVD using the analysis of heart sound. First, the presented meted applies spectral analysis for extracting time-varying spectral features which are able to effectively represent the complicated dynamics of the heart sound signals and later employ the strength of DL method for evaluating the estimated features for classification of the fundamental heart sounds.

Kumar and Mathur [12] proposed research of ML models for predicting CVD. The analysis of CVD employed 4 ML techniques like NB, SVM, KNN, and DT. Fan et al. [13] proved that the logistic regression method generated effective and accurate predictions for asymptomatic CAS amongst six ML algorithms. Demchenko and Kashirina [14] introduced a wide-ranging analysis of a CVD by using several methods. The main objective is to propose a powerful diagnostic model for this CVD. The real de-identified clinical data sets have been utilized for the model training, especially, samples of the MIMIC-III dataset. The usage of the Microsoft Azure ML platform inside this research enables us to design extremely scalable, reusable, and efficient classifier methods for CVD diagnosis.

Kigka et al. [15] proposed an ML method for the detection of higher risk plaques by employing non-invasive imaging and non-imaging-based aspects. Initially, implement a statistical analysis for identifying the important feature based on the results, and then, executed various FS methods and classification systems for developing ML approach. Terrada et al. [16] determined an MDSS of coronary heart ailment. This approach has the capacity of providing anticipations of heart diseases through the patients' healthcare records. The presented method is based on ML approaches like k-means and K-medoids clustering for the classification, ANN and KNN for predicting the absence and the presence of CVD.

3. The Proposed Model

In the current research, an effectual EDCA-FSSODL approach is presented for classifying and recognizing atherosclerosis. The proposed EDCA-FSSODL approach mainly aims to decrease the dimensionality of the features and diagnose atherosclerosis. The EDCA-FSSODL methodology comprises various operation phases like pre-processing, BCFPA based feature selection, BiLSTM based classification, and MRFO based hyperparameter tuning. Fig. 1 portrays the block diagram of EDCA-FSSODL methodology.
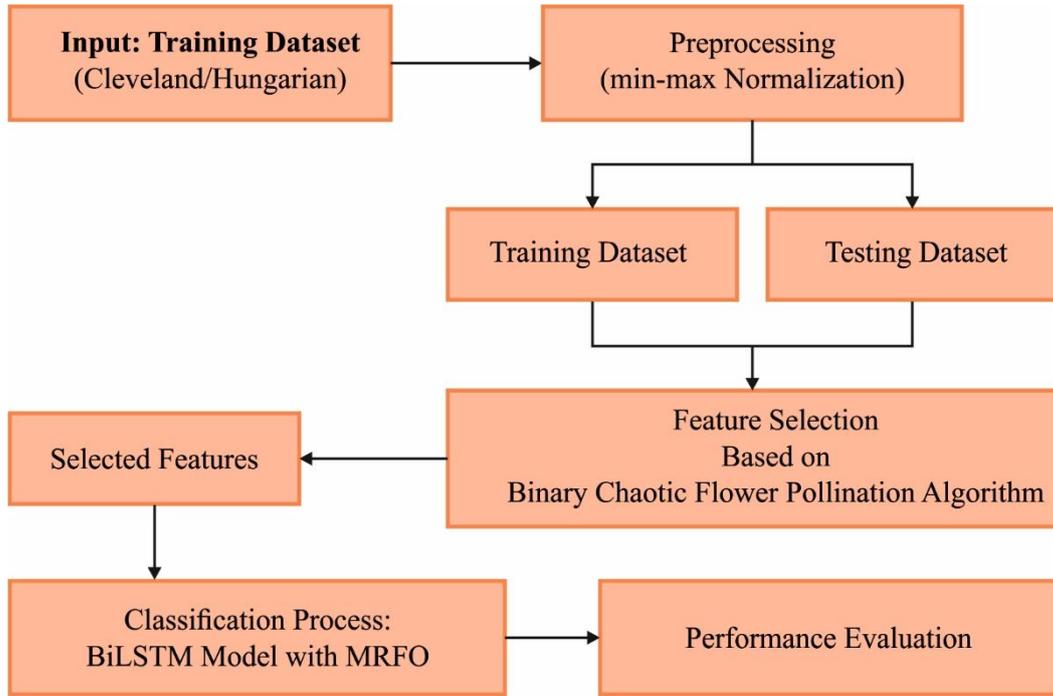
Fig. 1. Overall block presentation of EDCA-FSSODL approach

3.1 Level I: Data Pre-processing

ML methods are used to find out the tendency in the data set, by relative assessment amongst the dimension data point. When attempting to utilize ML method, an important problem is that their dimension has dramatically different scales. In this work, the min‐max normalized was employed for reducing the different scales of dimension. Normalized alters the information from a smaller range by performing linear transformation on original information [17]. The dimension value of the information is standardized within [0,1] by min‐max normalization.

$$t = \frac{v - \min_d}{\max_d - \min_d}(tran\_\max_d - tran\_\min_d) + tran\_\min_d \qquad (1)$$

whereas t represent the transformed value of v in d dimension, represent the original minimal value and max_d denotes the original maximal value of the d dimension. Likewise, tran_- min_d represent the transformed minimal value and tran_- max_d denotes the transformed maximal value of the d dimension.

3.2 Level II: Design of BCFPA Based Feature Selection Technique

Once the input data is preprocessed, the next stage is to design a new BCFPA technique to elect a feature subset. In 2012, Yang [18] suggested a nature enthused method known as FPA that is stimulated by the pollination method of flowering plans. The FPA is controlled by the four elementary rules below:

Abiotic and self‐pollinations are regarded as local pollination.

Biotic and cross‑pollinations are regarded as global pollination with pollen‑carrying pollinators carrying out LF.

Local together with global pollinations were governed by the switching possibility $p \in [0,1]$. Due to the physical proximity and another features namely wind, local pollination has an important fraction p from the complete pollination activity.

Flower constancy is assumed as the reproduction possibilities are proportionate to the similarity of 2 flowers.

In global pollination (Rules 1 and 3), the flower pollen is executed by pollinators like insects. The pollen travels a longer distance by means of Lévy flight (LF) distribution. It is mathematically expressed by,

$$x_i^{(t+1)} = x_i^t + \alpha L(\lambda)(g^* - x_i^t) \tag{2}$$

Whereas,

$$L(\lambda) = \frac{\lambda . \Gamma(\lambda) . \sin(\lambda)}{\pi} . \frac{1}{s^{1+\lambda}}, s > 0, \tag{3}$$

Here, x_i^t indicates the pollen i at the iteration t and g^* indicates the present optimal solution amongst each solution at the present generation, whereas α denotes the scaling feature for controlling the size of the step, s, L(λ) signifies the step and LF step size respective to the pollination strength and Γ(λ) indicates the gamma function, while the λ value is in 1≤λ≤2.

The Local Pollination (Rule 2) is scientifically expressed by [19],

$$x_i^{(t+1)} = x_i^t + \varepsilon\left(x_j^t - x_k^t\right), \tag{4}$$

Let x_j^t and x_k^t be the pollen from dissimilar flowers j and k of alike plant species. To stimulate the global and local flower pollinations, the switching probability P was employed (Rule 4).

The CFPA technique is derived by the integration of the concepts of FPA with chaos theory. Chaos, as an extensive non-linear phenomenon in nature, has the features of sensitivity, randomness, ergodicity to key condition, etc. Due to the features of arbitrariness and ergodicity, chaotic motion could be traversing every state in some range as per the own laws without reiteration. Hence, chaos variables are utilized to search maximally, the study would certainly have more merits when associated to random search. Chaotic sequence created by distinct mappings like tent, singer, logistic, sine, and sinusoidal map. In the study, numerous chaotic maps have been attempted and the maximal one is elected for joining the BFO method [20]. As per the study, logistic map has obtained the optimal outputs.

$$x_{i+1} = u x_i (1 - x_i) \tag{5}$$

u denotes the control variable and let u=4. If u=4, the logistic mapping comes into an elaborated chaotic state. Assume x_i∈(0,1) and x_i≠0.25,0.5,0.75.

The primary bacterial populace θ is mapped to the chaotic serial that is formulated as per (5), leading to a chaotic bacterial population pch.

$$pch = x_i * \theta \tag{6}$$

BCFPA is presented by [21], where the searching space is built as a d-dimension Boolean lattice, where the solution is upgraded over the corners of hypercube. While FS problem has to choose a certain feature or not, hence the solutions are denoted as binary vectors, whereas 1 shows a feature would be selected for comprising the novel data set and 0 if not. Sigmoid function is implemented for building binary vector as:

$$\left(x_i^j(t)\right) = \frac{1}{1 + e^{-x_i^j(t)}}, \tag{7}$$

$$x_i^j(t) = \begin{cases} 1 & if \ S\left(x_i^j(t)\right) > \sigma, \\ 0 & otherwise, \end{cases} \tag{8}$$

Whereas $x_i^j(t)$ signifies the novel pollen (solution) i using j^{th} feature vector, while $i = 1,2,.....m$ and $j = 1,2,.....d$, at the iteration t and $\sigma \sim U(0,1)$.

3.3 Level III: Data Classification by employing BiLSTM Approach

During data classification procedure, the chosen features are applied to the BiLSTM classifier. New improvements from massive data accessibility and DL approaches are developing the time series data driven forecast further affordable and prestigious containing predicting and vision associated tasks like crowd examination [22]. But the recognized works on LSTM to time sequence issues provide a vital direction to SOH evaluation of battery and power utilization. The LSTM network is a kind of RNN, in which is extremely well-known to process serial data and address the long-term memory disadvantages of vanilla RNN. The LSTM extended the framework of RNN utilizing a gate process and distinct memory cell that regulate the flow of data throughout the networks. Fig. 2 portrays the framework of Bi-LSTM method.
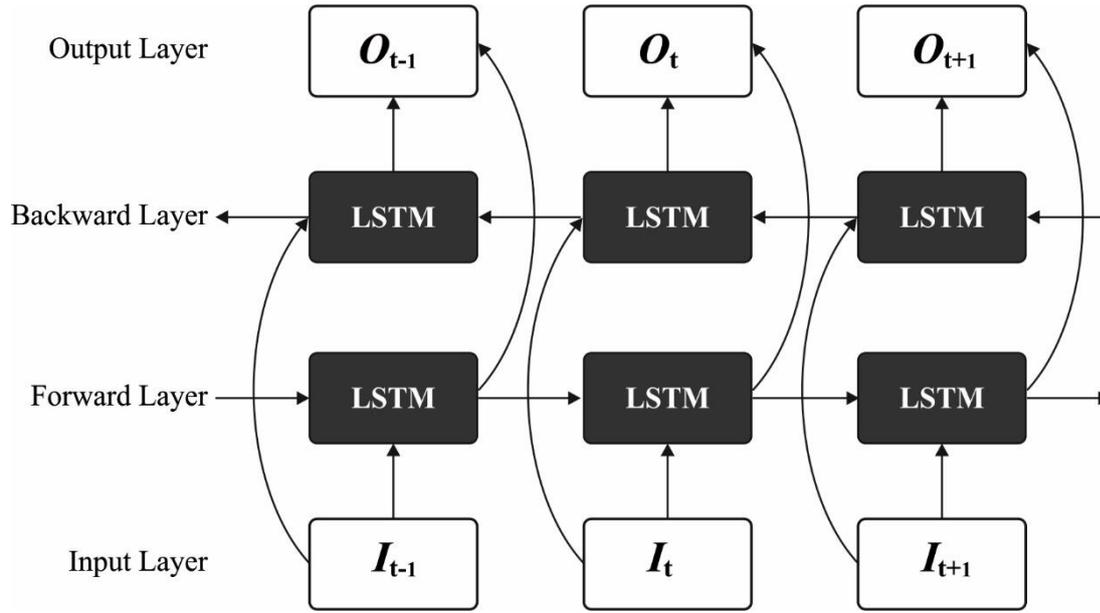
Fig. 2. Structure of Bi-LSTM

The gate model contains 3 units namely input, forgot, and output gate. These gates manage the flow of data through the network for allowing that data requires for persisting or as long as it can persist then reading it under the cell memory. The LSTM network was able for retaining the key data and remove the lesser essential data. The memory cell is a recurring self-related unit called a constant error carousel (CEC) that is a state vector for preserving long-term dependability. Afterward, for distinguishing self-contained cell memory in convention state h_t from LSTM, it can be signified as c_t. The forget gate f_t attains input x_t and h_{t-1} for determining that data requires that recollected from c_{t-1}. The activation function to gates i_τ, o_t, and f_t are sigmoid layers in which all values are presented amongst zero and one, but c_{t-1} offers the data retentions for describing the scale. More details were away from the scope of work. But the above-mentioned procedure was formally determined in Eq. (9)-(13):

$$i_t = \sigma(W_{xi}x_\tau + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \quad (9)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \quad (10)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_\tau + b_0) \quad (11)$$

$$c_t = f_\tau \circ c_{t-1} + i_t \circ \emptyset(W_{xc}x_\tau + W_{hc}h_{t-1} + b_c) \quad (12)$$

$$h_t = 0_\tau \circ \emptyset(c_t) \quad (13)$$

At this point, W_{xi}, W_{hi}, W_{ci}, W_{xf}, W_{hf}, W_{cf} W_{xo}, W_{ho}, W_{co}, W_{xc}, and W_{hc} represents the weight matrices to gates and cell memory states, but h_{t-1} refers the preceding hidden state, and c_t stands for the cell state. The bias of gate was signified with b_i, b_o, b_f, and b_c but 0 referring the element-wise multiplication process. At the same time, σ, and $\emptyset$ illustrate the logistic sigmoid and hyperbolic tangent function. This activation function is determined as individually in Eq. (14) and (15):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{14}$$

$$\emptyset(x) = \frac{e^x - e^{-x}}{e^X + e^{-x}} \tag{15}$$

The BiLSTM has 2 parallel LSTM layers like forwarding and backward directions. As the input was processed twice, BiLSTMs remove supplementary input data. Thus, improving contextual data to make optimum forecasts than LSTM. So, the BiLSTMs offer faster convergence and accuracy than LSTMs.

The resultant of all LSTM is integrated based on the subsequent formula:

$$y_t = W_{\vec{h}y}\overrightarrow{h_t} + W_{\overleftarrow{h}y}\overleftarrow{h_t} + b_y, \tag{16}$$

Where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ refers the resultants of forward together with backward LSTM.

3.4 Level IV: Hyperparameter Tuning by employing MRFO Approach

For maximal tuning process of BiLSTM method, the MRFO approach is used. MRFO model was developed by Zhao et al. [23]. It stimulates the Manta ray behaviors. MRFO is expressed with duplicating the hunting approaches i.e., somersault, chaining, and cyclone. Although MRFO is analogous to various metaheuristic approaches, its initialization phase is determined as follows.

$$X_{ij}(:) = Lb_{ij} + r(:) \cdot (Hb_{ij} - Lb_{ij}) \forall i \in N_{pop}, j \in N_{var} \tag{17}$$

The Manta ray accomplishment was altered to the optimal food source in the chaining strategy as follows

$$X_{i,j}(t+1) = \begin{cases} X_{i,j}(t) + \left(X_{best,j}(t) - X_{i,j}(t)\right) \cdot (r(:) - \lambda)\Lambda\forall i = 1, j \in N_{var} \\ X_{i,j}(t) + \left(X_{i-1,j}(t) - X_{i,j}(t)\right) \cdot (r(:) - \lambda)\Lambda\forall i > 1: \ N_{pop} \end{cases} \tag{18}$$

$$\phi = 2 \cdot r(:) \cdot \sqrt{|\log(r(:))|} \tag{19}$$

At the same time, swarm will swim in a spiral analogous to cyclones afterward observing the food source. Finally, the modern Manta ray position is modified surrounding the optimal location in this operating system as follows.

$$X_{i,j}(t+1) = X_{i,j}(t) + SF \cdot \left(r_2 X_{best,j} - r_3 X_{i,j}(t)\right) \Lambda\forall i \in N_{pop} \tag{20}$$

X_{best}, j is determined as shown in (17) that describes optimal position with high [24]. When the lower concentration is confirmed, the arbitrary location is set as illustrated in (24).

$$X_{i,j}(t+1) = \begin{cases} X_{best,j} + \Lambda X_{i,j} \cdot (r - \beta)\Lambda\forall i = 1, j \in N_{va} \\ X_{best,j} + r \cdot (x_{i-1,j}(t) - X_{i,j}(t)) + \beta \cdot \Lambda X_{i,j}\Lambda\forall i > 1: N_{pop} \end{cases}$$

$$where \ \Lambda X_{i,j} = X_{best,j}(t) - X_{i,j}(t) \tag{21}$$

$$\beta = 2 \cdot \exp\left(r_1 \frac{T_{\max} - t + 1}{T_{\max}}\right) \cdot sin(2\pi r_1) \tag{22}$$

$$X_{i,j}(t+1) = \begin{cases} X_{rand} + \Lambda X_{i,j}.(r - \beta)\Lambda \forall i = 1, j \in N_{var} \\ X_{best,j} + r.(X_{i-1,j}(t) - X_{i,j}(t)) + \beta \cdot \Lambda X_{i,j}\Lambda \forall i > 1: N_{pop} \end{cases}$$

$$where \;\; \Lambda X_{i,j} = X_{rand} - X_{i,j}(t) \tag{23}$$

$$X_{rand}(:) = Lb + r(:) \cdot (Hb - Lb) \tag{24}$$

Indeed, meta heuristic algorithm shares two certain behaviors exploitation and exploration. According to this behavior, meta-heuristic algorithms have superiority over other methods. In MRFO, when $t/T_{\max}$ is lesser than rand, the cycle of exploitation is carried out, or else search space would be executed. N_{pop}, $T_{\max}$, and SF describes the MRFO controlling scheme, that must be followed sensibly to guarantee its better performances. The MRFO technique grows an FF to attain improved classification performance. It states a positive number for portraying the optimal effectualness of the candidate solution. In this case, the minimization of the classification rate of error was regarded as FF, as illustrated in Eq. (25). Optimal and worst solutions attain minimal and maximal rate of error.

$$fitness(x_i) = ClassifierErrorRate(x_i)$$

$$= \frac{numerb \; of \; misclassified \; instances}{Total \; number \; of \; instances} * 100 \tag{25}$$

4. Performance Validation

This process of the EDCA-FSSODL technique occurs utilizing Python 3.6.5 tool. The outcomes are investigated by implementing two standard datasets namely Cleveland and Hungarian datasets [25]. The first Cleveland dataset includes a set of 76 features, and 14 features are generally utilized. The class distribution of the Cleveland dataset is demonstrated in Fig. 3, indicating 164 instances under 164 healthy classes and 139 instances under abnormal classes.
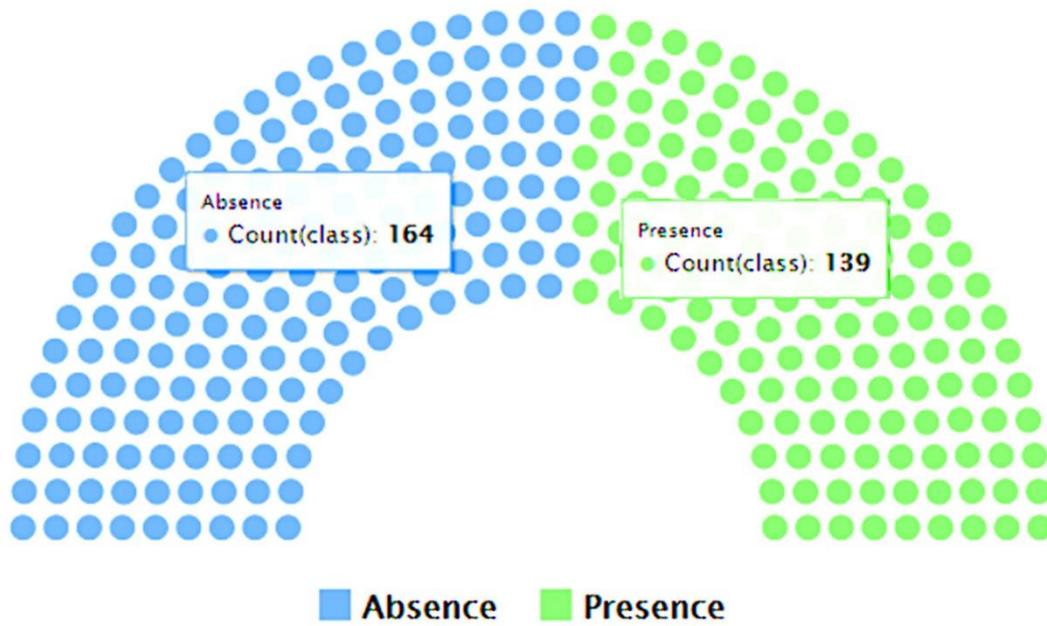
Fig. 3. Class distribution of Cleveland Dataset

The next Hungarian dataset comprises a total of 14 features. The class distribution of the Cleveland dataset is portrayed in Fig. 4, indicating 188 instances under 106 healthy classes and 139 instances under abnormal classes.
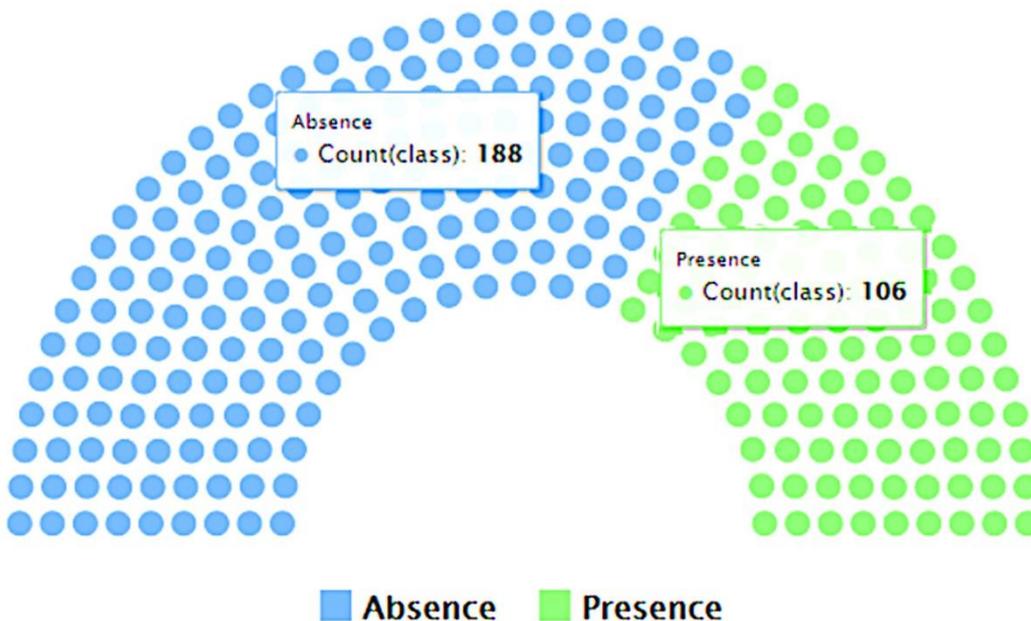
Fig. 4. Class distribution of Hungarian Dataset

The feature selection results of the presented model are portrayed in Table 1. The table denotes that the presented model has selected a set of 7 and 6 factors from the datasets of test Cleveland and Hungarian respectively.

<div style="text-align:center">**Table 1** FS of proposed method with 2 dataset</div>

Datasets	Selected Numbers	Selected features
Cleveland	07	Age, trestbps, Restecg, Exang, Oldpeak, slope, Thal
Hungarian	06	Age, trestbps, chol, Thalach, Exang, Oldpeak
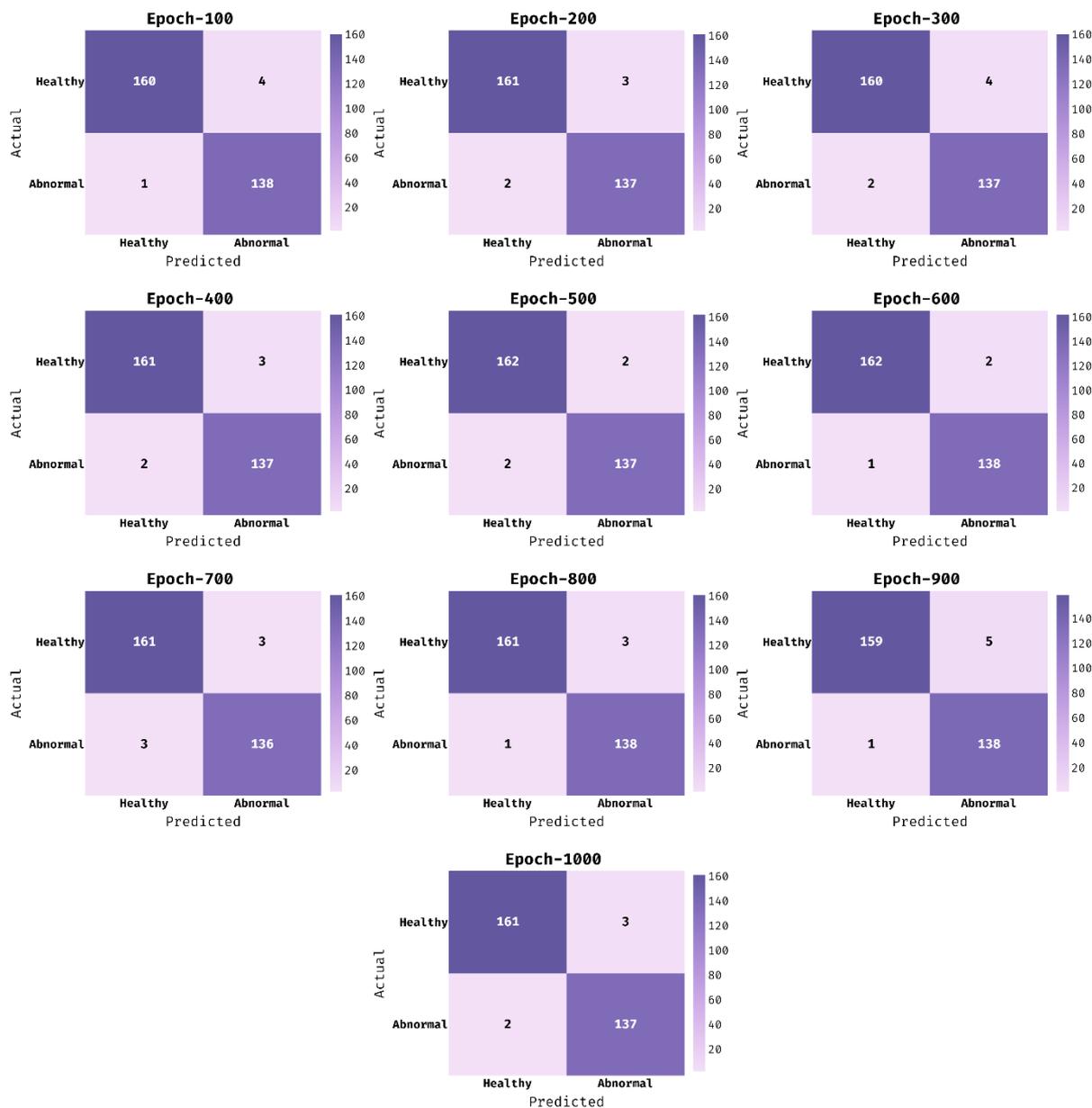
Fig. 5. Confusion matrix of EDCA-FSSODL method on Cleveland dataset

Fig. 5 demonstrates the confusion matrices set generated by the EDCA-FSSODL approach on the testing Cleveland dataset under distinct epochs. As a sample, with 100 epochs, the EDCA-FSSODL approach has classified 160, and 138 instances into Healthy and Abnormal class. Additionally, with 300 epochs, the EDCA-FSSODL technique

has classified 160, and 137 instances into Healthy and Abnormal class. Moreover, with 700 epochs, the EDCA-FSSODL technique has classified 161, and 136 instances into Healthy and Abnormal class. Furthermore, with 1000 epochs, the EDCA-FSSODL approach has classified 161, and 137 instances into Healthy and Abnormal class.

A brief classification output evaluation of the EDCA-FSSODL technique on the dataset of the test Cleveland is given in Table 2 and Fig. 6. The investigational outcomes portrayed that the EDCA-FSSODL technique has achieved effective classification results under all epochs. As a sample, on 100 epochs, the EDCA-FSSODL methodology has offered $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 97.56%, 99.28%, 98.35%, 98.46%, and 96.70%. Simultaneously, on 400 epochs, the EDCA-FSSODL methodology has offered $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.17%, 98.56%, 98.35%, 98.47%, and 96.68%. Concurrently, on 700 epochs, the EDCA-FSSODL method has existing $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.17%, 97.84%, 98.02%, 98.17%, and 96.01%. Lastly, on 1000 epochs, the EDCA-FSSODL method has presented $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.17%, 98.56%, 98.35%, 98.47%, and 96.68%.

Table 2 Result evaluation of EDCA-FSSODL method on Cleveland dataset

Epoch Numbers	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}	MCC
Epoch-100	97.56	99.28	98.35	98.46	96.70
Epoch-200	98.17	98.56	98.35	98.47	96.68
Epoch-300	97.56	98.56	98.02	98.16	96.03
Epoch-400	98.17	98.56	98.35	98.47	96.68
Epoch-500	98.78	98.56	98.68	98.78	97.34
Epoch-600	98.78	99.28	99.01	99.08	98.01
Epoch-700	98.17	97.84	98.02	98.17	96.01
Epoch-800	98.17	99.28	98.68	98.77	97.35
Epoch-900	96.95	99.28	98.02	98.15	96.05
Epoch-1000	98.17	98.56	98.35	98.47	96.68
Average	**98.05**	**98.78**	**98.38**	**98.50**	**96.75**
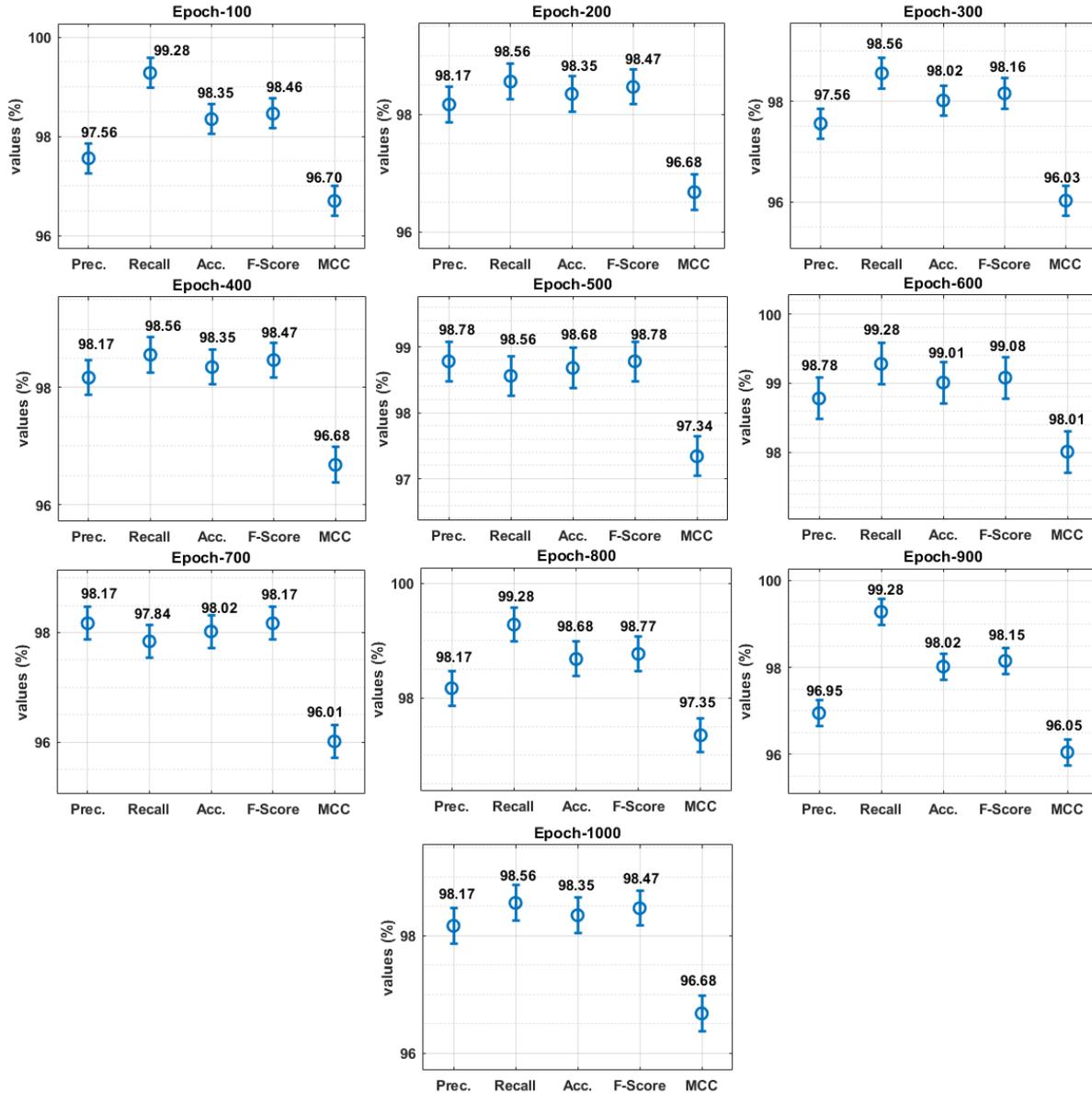
Fig. 6. Result evaluation of EDCA-FSSODL method on Cleveland dataset

Fig. 7 exhibits the accuracy assessment of the EDCA-FSSODL methodology on the test Cleveland dataset. The outputs demonstrated that the EDCA-FSSODL methodology has achieved enhanced accomplishment with augmented accuracy values of training and validation. It is seen that the EDCA-FSSODL technique attained higher accuracy of validation over training.

Fig. 8 exhibits the loss assessment of the EDCA-FSSODL approach on the test Cleveland dataset. The results demonstrated that the EDCA-FSSODL approach has accomplished enhanced performance with reduced loss values of training and validation. It is seen that the EDCA-FSSODL technique attained higher loss values of validation over training.
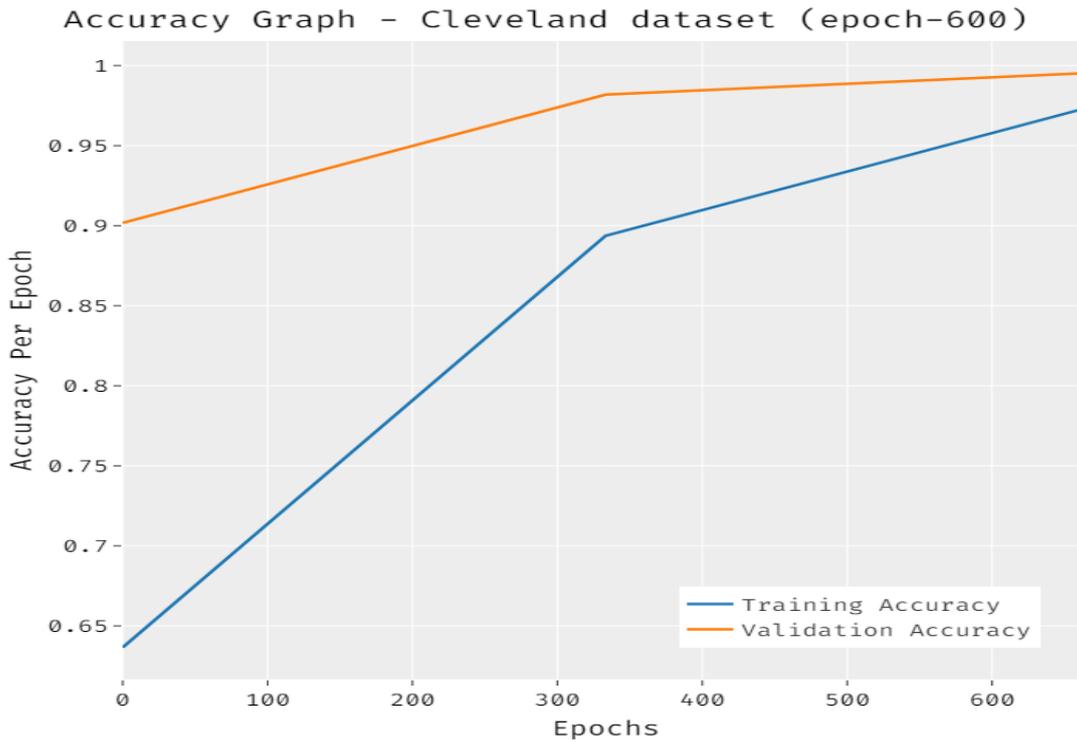
Fig. 7. Accuracy evaluation of EDCA-FSSODL method on Cleveland dataset
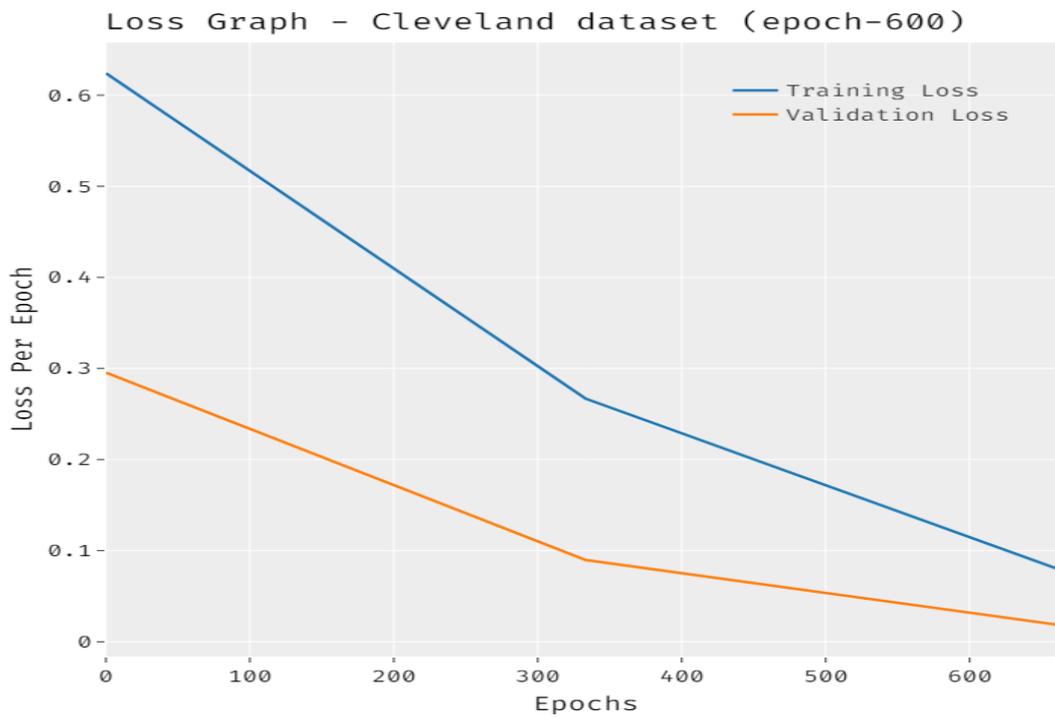
Fig. 8. Loss graph evaluation of EDCA-FSSODL method on Cleveland dataset

Finally, the comparative result evaluation of the EDCA-FSSODL approach is made with current models on the Cleveland dataset and is portrayed in Table 3 and Fig. 9. The outputs exhibit that the DT and RF methods have resulted to lower classifying accomplishment over the other methods. It is seen that the NB and LOGR models have accomplished somewhat improved classifier results.

Table 3 Comparative analysis of EDCA-FSSODL method on Cleveland dataset

Approaches	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}
Decision Tree	57.10	66.70	62.80	61.50
GBT Model	59.50	73.30	74.70	65.70
LOGR Algorithm	60.00	77.80	73.30	67.70
MLP-CA	74.20	65.70	74.00	69.70
Naïve Bayes	67.90	55.90	70.70	61.30
Random Forest	77.10	58.70	67.90	66.70
EDCA-FSSODL (WOFS)	86.81	81.26	84.49	84.93
EDCA-FSSODL (WFS)	98.05	98.78	98.38	98.50
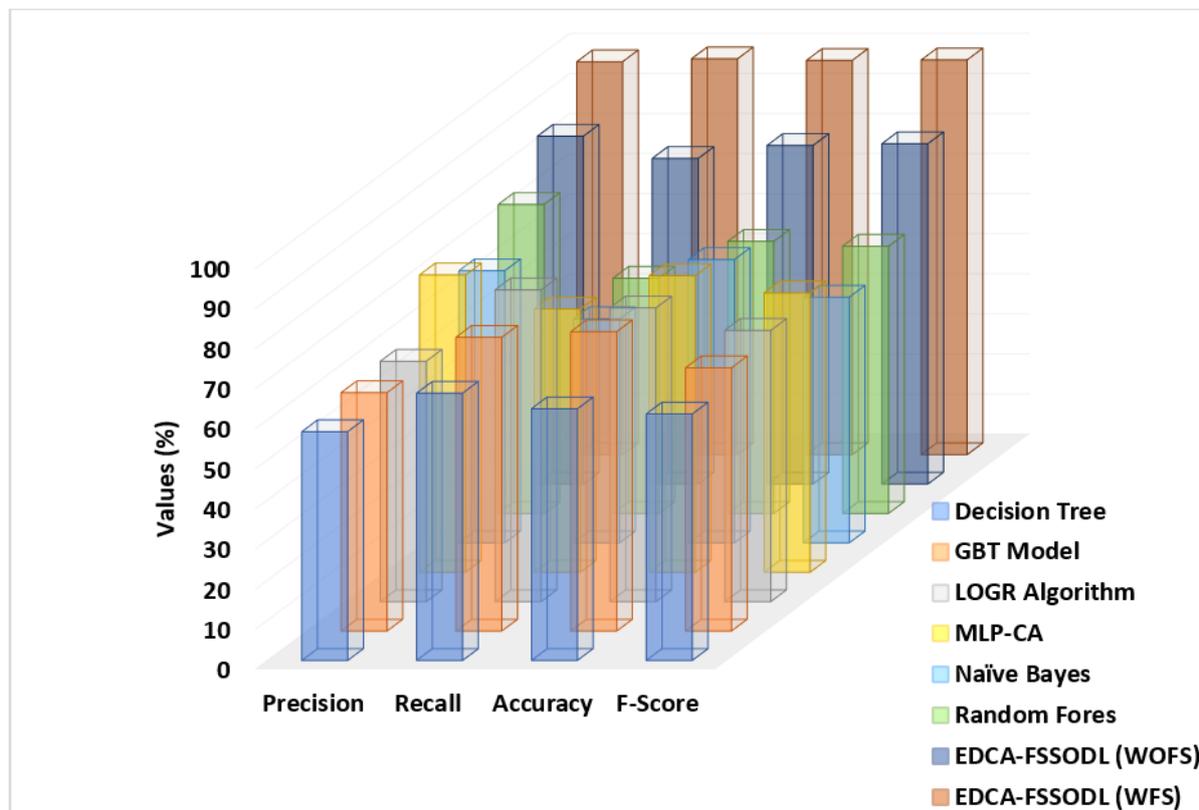
Fig. 9. Comparative analysis of EDCA-FSSODL method on Cleveland dataset

Moreover, the MLP-CA and GBT models have reached moderately closer classification performance. However, the EDCA-FSSODL (WFS) technique has reached superior classification performance with the $prec_n$, $reca_l$, $accu_y$, and F_{score} of 98.05%, 98.78%m 98.38%, and 98.50% respectively.
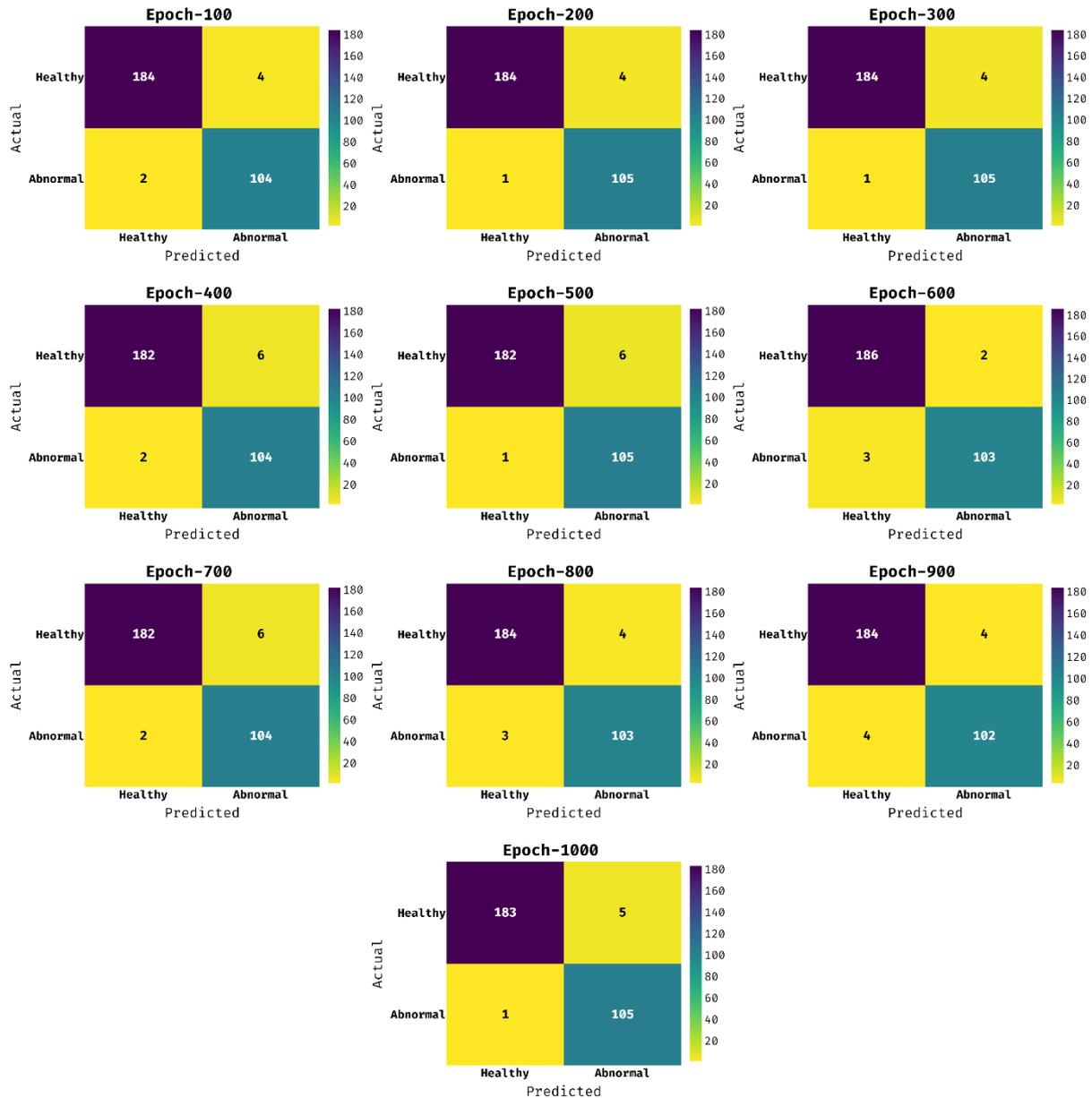
Fig. 10. Confusion matrix analysis of EDCA-FSSODL method on Hungarian dataset

Fig. 10 reveals the confusion matrix set generated by the EDCA-FSSODL method on the test Hungarian dataset under changing epochs. As a sample, with 100 epochs, the EDCA-FSSODL approach has classified 184, and 104 instances into Healthy and Abnormal class. Besides, with 300 epochs, the EDCA-FSSODL methodology has classified 184, and 105 instances into Healthy and Abnormal class. In addition, with 700 epochs, the EDCA-FSSODL technique has classified 182, and 104 instances into Healthy and Abnormal class. Eventually, with 1000 epochs, the EDCA-FSSODL technique has classified 183, and 105 instances into Healthy and Abnormal class.

A detailed classification output evaluation of the EDCA-FSSODL methodology on the test Hungarian dataset is given in Table 4 and Fig. 11. The investigational outcomes portrayed that the EDCA-FSSODL methodology has achieved effectual classification outputs under all epochs. As a sample, on 100 epochs, the EDCA-FSSODL technique has offered $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.92%, 97.87%, 97.96%, 98.40%, and 95.60%. Simultaneously, on 400 epochs, the EDCA-FSSODL approach has existing $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.91%, 96.81%, 97.28%, 97.85%, and 94.19%. Then, on 700 epochs, the EDCA-FSSODL technique has offered $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 98.91%, 96.81%, 97.28%, 97.85%, and 94.19%. At last, on 1000 epochs, the EDCA-FSSODL approach has offered $prec_n$, $reca_l$, $accu_y$, F_{score}, and MCC of 99.46%, 97.34%, 97.96%, 98.39%, and 95.65%.

Table 4 Result evaluation of EDCA-FSSODL method on Hungarian dataset

No. of Epochs	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}	MCC
Epoch-100	98.92	97.87	97.96	98.40	95.60
Epoch-200	99.46	97.87	98.30	98.66	96.36
Epoch-300	99.46	97.87	98.30	98.66	96.36
Epoch-400	98.91	96.81	97.28	97.85	94.19
Epoch-500	99.45	96.81	97.62	98.11	94.95
Epoch-600	98.41	98.94	98.30	98.67	96.31
Epoch-700	98.91	96.81	97.28	97.85	94.19
Epoch-800	98.40	97.87	97.62	98.13	94.85
Epoch-900	97.87	97.87	97.28	97.87	94.10
Epoch-1000	99.46	97.34	97.96	98.39	95.65
Average	**98.93**	**97.61**	**97.79**	**98.26**	**95.26**
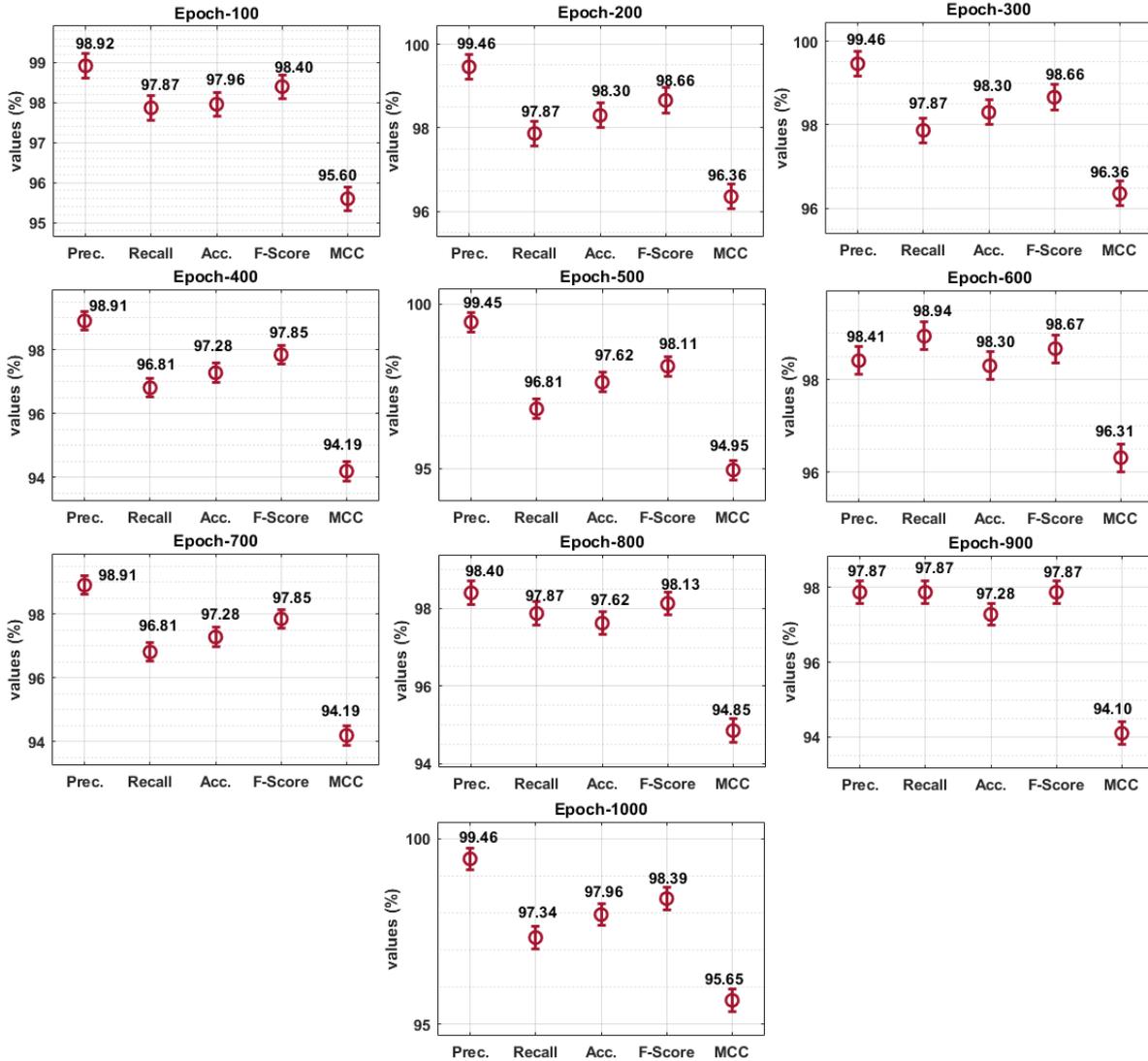
Fig. 11. Result evaluation of EDCA-FSSODL method on Hungarian dataset

Fig. 12 illustrates the accuracy evaluation of the EDCA-FSSODL methodology on the test Hungarian dataset. The outputs demonstrated that the EDCA-FSSODL methodology has achieved enhanced accomplishment with maximum accuracy values of training and validation. It can be clear that the EDCA-FSSODL technique attained enhanced accuracy of validation over training.

Fig. 13 illustrates the loss evaluation of the EDCA-FSSODL method on the test Hungarian dataset. The outputs demonstrated that the EDCA-FSSODL method has given an output in an enhanced output with the mitigated loss values of training and validation. It can be clear that the EDCA-FSSODL approach attained decreased loss of validation over training.
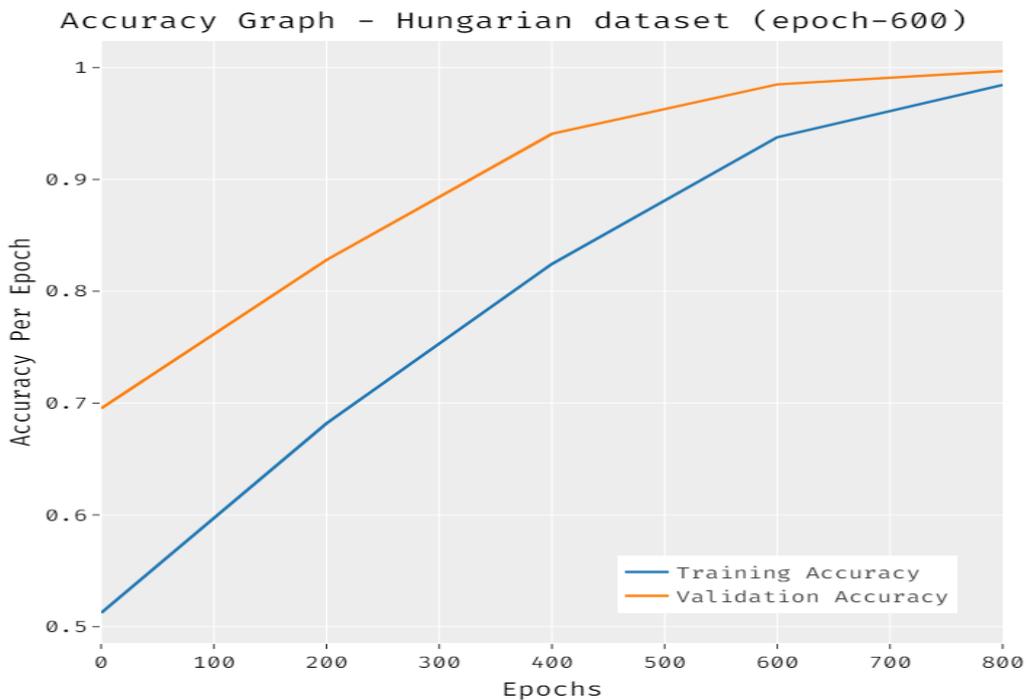
Fig. 12. Accuracy evaluation of EDCA-FSSODL method on Hungarian dataset
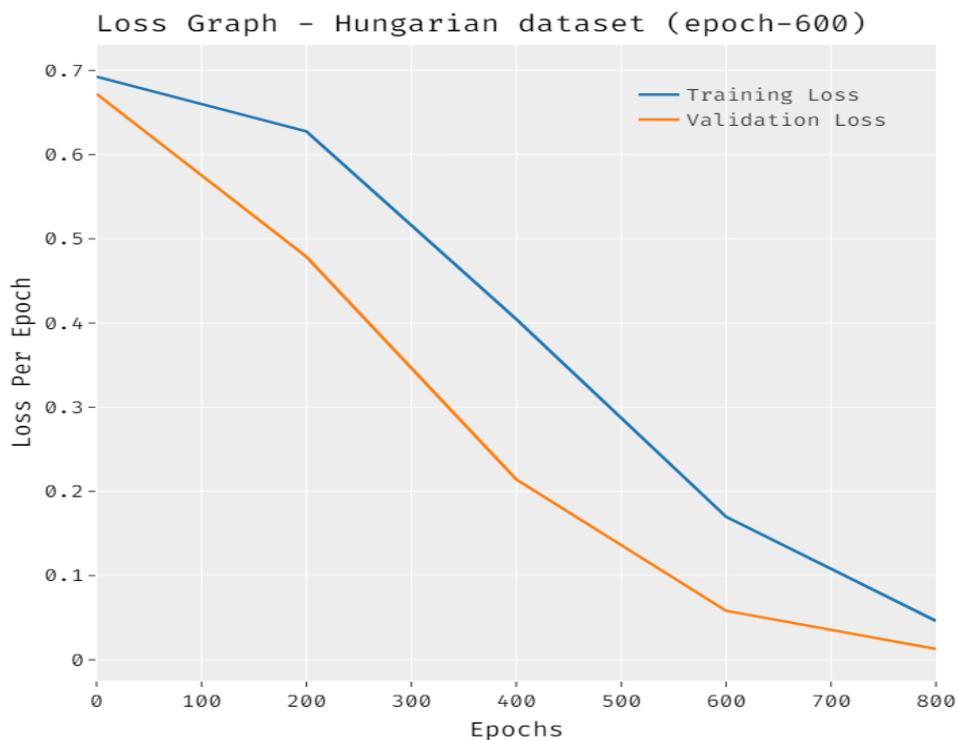
Fig. 13. Loss graph evaluation of EDCA-FSSODL method on Hungarian dataset

Table 5 Relative evaluation of EDCA-FSSODL method on Hungarian dataset

Methods	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}
Decision Tree	88.00	78.60	88.80	83.00
GBT Model	87.00	80.00	89.70	83.30
LOGR Algorithm	91.40	94.10	94.90	92.80
MLP-CA	92.60	92.60	95.50	92.60
Naïve Bayes	63.60	72.40	78.00	67.70
Random Forest	93.10	87.10	93.20	90.00
EDCA-FSSODL (WOFS)	96.04	97.02	96.34	96.44
EDCA-FSSODL (WFS)	98.93	97.61	97.79	98.26

Lastly, the relative outcome evaluation of the EDCA-FSSODL methodology is made with current approaches on the Hungarian dataset is portrayed in Table 5 and Fig. 14 [26]. The outputs portrayed that the DT and RF methods have resulted to lower classification accomplishment over the other methods. Also, the NB and LOGR systems have accomplished slightly increased classifier results. Besides, the MLP-CA and GBT models have obtained to moderately closer classification performance.
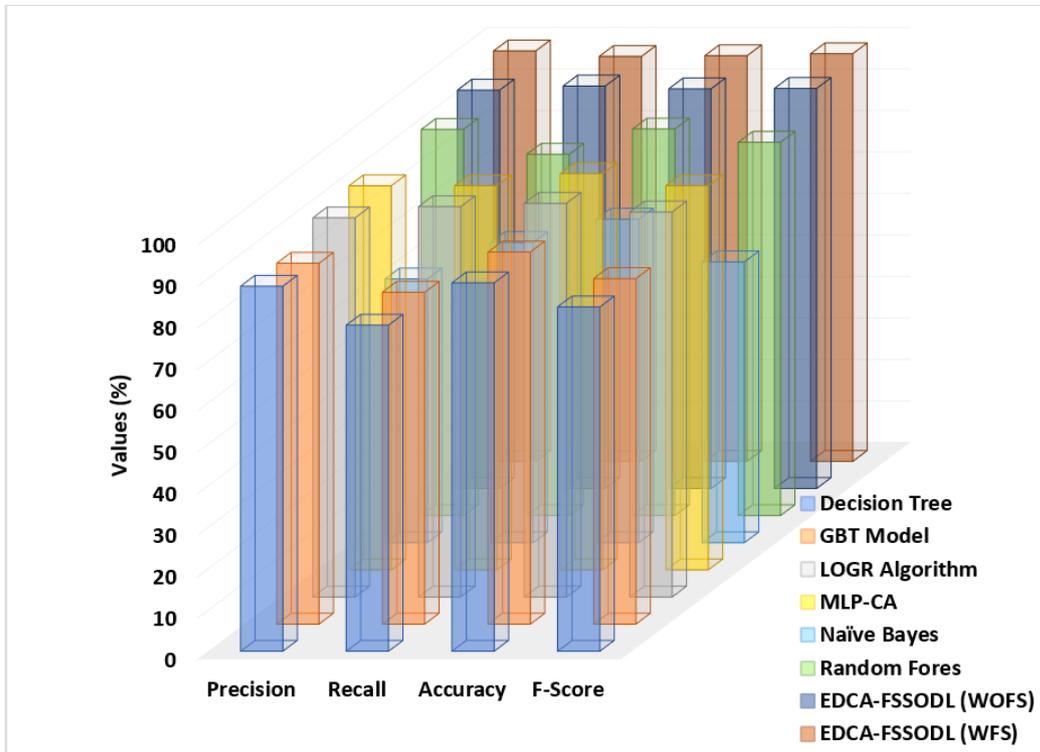
Fig. 14. Comparative analysis of EDCA-FSSODL technique on Hungarian dataset

But the EDCA-FSSODL (WFS) algorithm has attained superior classification performance with the $prec_n$, $reca_l$, $accu_y$, and F_{score} of 98.93%, 97.61%m 97.79%, and 98.26% correspondingly. By looking into the detailed experimental result evaluation, it is ensured that the EDCA-FSSODL technique has the capacity to detect and classify heart disease effectually using biomedical data.

5. Conclusion

In this study, an effectual EDCA-FSSODL methodology is presented for the identification and classification of atherosclerosis. The suggested EDCA-FSSODL methodology mainly aims to decrease the dimensionality of the features and diagnose atherosclerosis. The EDCA-FSSODL technique comprises various operational phases like pre-processing, BCFPA based feature selection, BiLSTM based classification, and MRFO based hyperparameter tuning. The design of BCFPA based feature selection and MFRO based hyperparameter selection process helps to considerably augment the classifying results. For investigating the improvements of the EDCA-FSSODL technique, a wide range of investigational analyses is performed against two standard datasets and validate the outputs by means of diverse measures. A comprehensive relative study highlighted the supremacy of the EDCA-FSSODL technique over the other approaches.

References

[1] Terrada, O., Cherradi, B., Raihani, A. and Bouattane, O., 2020. A novel medical diagnosis support system for predicting patients with atherosclerosis diseases. Informatics in Medicine Unlocked, 21, p.100483.

[2] V.S.H. Rao, M.N. Kumar, Novel approaches for predicting risk factors of atherosclerosis, IEEE J. Biomed. Health Inform. 17 (2012) 183–189.

[3] J.J. Batzel, M. Bachar, F. Kappel, Mathematical Modeling and Validation in Physiology: Applications to the Cardiovascular and Respiratory Systems, vol. 2064, Springer, 2012.

[4] B. Iba nez, J.J. Badimon, M.J. Garcia, Diagnosis of atherosclerosis by imaging, Am. J. Med. 122 (2009) S15–S25.

[5] Daanouni O, Cherradi B, Tmiri A. Type 2 diabetes mellitus prediction model based on machine learning approach. In: Ben Ahmed M, Boudhir AA, Santos D, El Aroussi M, Karas ̇ IR, editors. In Innovations in smart cities applications edition 3. Cham: Springer International Publishing; 2020. p. 454–69.

[6] Abdar M, Acharya UR, Sarrafzadegan N, Makarenkov V. NE-nu-SVC: a new nested ensemble clinical decision support system for effective diagnosis of coronary artery disease. IEEE Access 2019;7:167605–20.

[7] Terrada O, Cherradi B, Raihani A, Bouattane O. A fuzzy medical diagnostic support system for cardiovascular diseases diagnosis using risk factors. In: 2018 international Conference on electronics, control, Optimization and computer science (ICECOCS); Dec. 2018. p. 1–6.

[8] Nasarian E, et al. Association between work-related features and coronary artery disease: a heterogeneous hybrid feature selection integrated with balancing approach. Pattern Recogn Lett May 2020;133:33–40.

[9] Y. Hu, L. Li, L. Shen, H. Gao, The relationship between arterial wall stiffness and left ventricular dysfunction, Neth. Heart J. 21 (2013) 222–227.

[10]   Jain, K., Jain, S., Guha, A. and Patra, A., 2021. An approach to early stage detection of atherosclerosis using arterial blood pressure measurements. Biomedical Signal Processing and Control, 68, p.102594.

[11]   Upretee, P. and Yüksel, M.E., 2021. Accurate classification of heart sounds for disease diagnosis by using spectral analysis and deep learning methods. In Data Analytics in Biomedical Engineering and Healthcare (pp. 215-232). Academic Press.

[12]   Kumar, B. and Mathur, H., 2021. Comprehensive Analysis of Atherosclerosis Disease Prediction using Machine Learning. Annals of the Romanian Society for Cell Biology, pp.17962-17975.

[13]   Fan, J., Chen, M., Luo, J., Yang, S., Shi, J., Yao, Q., Zhang, X., Du, S., Qu, H., Cheng, Y. and Ma, S., 2021. The prediction of asymptomatic carotid atherosclerosis with electronic health records: a comparative study of six machine learning models. BMC medical informatics and decision making, 21(1), pp.1-9.

[14]   Demchenko, M. and Kashirina, I., 2020. The Use of Machine Learning Methods to the Automated Atherosclerosis Diagnostic and Treatment System Development.

[15]   Kigka, V.I., Sakellarios, A.I., Mantzaris, M.D., Tsakanikas, V.D., Potsika, V.T., Palombo, D., Montecucco, F. and Fotiadis, D.I., 2021, November. A Machine Learning Model for the Identification of High risk Carotid Atherosclerotic Plaques. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 2266-2269). IEEE.

[16]   Terrada, O., Cherradi, B., Raihani, A. and Bouattane, O., 2019, April. Classification and Prediction of atherosclerosis diseases using machine learning algorithms. In 2019 5th International Conference on Optimization and Applications (ICOA) (pp. 1-5). IEEE.

[17]   Khare, N., Devan, P., Chowdhary, C.L., Bhattacharya, S., Singh, G., Singh, S. and Yoon, B., 2020. Smo-dnn: Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection. Electronics, 9(4), p.692.

[18]   Yang, X.S., 2012, September. Flower pollination algorithm for global optimization. In International conference on unconventional computing and natural computation (pp. 240-249). Springer, Berlin, Heidelberg.

[19]   Nguyen, T.T., Pan, J.S. and Dao, T.K., 2019. An improved flower pollination algorithm for optimizing layouts of nodes in wireless sensor network. Ieee Access, 7, pp.75985-75998.

[20]   Cai, Z., Gu, J., Wen, C., Zhao, D., Huang, C., Huang, H., Tong, C., Li, J. and Chen, H., 2018. An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach. Computational and mathematical methods in medicine, 2018.

[21]   Rodrigues, D., Yang, X.-S., Souza, A. N., and Papa, J. P. (2015). Binary flower pollination algorithm and its application to feature selection. In Yang, X.-S., editor, Recent Advances in Swarm Intelligence and Evolutionary Computation, volume 585 of Studies in Computational Intelligence, pages 85–100. Springer International Publishing.

[22]   Khan, N.; Ullah, A.;Haq, I.U.; Menon, V.G.; Baik, S.W.; SD-Net: Understanding overcrowded scenes in real-time via an efficient dilated convolutional neural network. J. Real-Time Image Process. 2021, 18, 1729–1743.

[23]   Zhao, W., Zhang, Z., Wang, L., 2020. Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications. Eng. Appl. Artif. Intell. 87, 103300.

[24]   Hassan, M.H., Houssein, E.H., Mahdy, M.A. and Kamel, S., 2021. An improved manta ray foraging optimizer for cost-effective emission dispatch problems. Engineering Applications of Artificial Intelligence, 100, p.104155.

[25]   https://archive.ics.uci.edu/ml/datasets.php

[26]   Gárate-Escamila, A.K., El Hassani, A.H. and Andrès, E., 2020. Classification models for heart disease prediction using feature selection and PCA. Informatics in Medicine Unlocked, 19, p.100330.