

A Deep Learning Approach with Binary Particle Swarm Optimization for Optimizing Prediction of Heart Disease

Dr. K. Sampathkumar¹ and S. Periyasamy²

¹ Guest Lecturer, Department of Mathematics, Government Arts College (Grade-1), Ariyalur, India

² Guest Lecturer, Department of Computer Science, Government Arts College, Harur, India

E-mail: sampathmath81@gmail.com

Received 28-02-2024

Accepted for publication 15-03-2024

Published 20-03-2024

Abstract

The significant component of the body part is heart function leads to human life prediction. Heart failure may be inactive, but it can be predicted through health screenings and the individual's daily activities. The contemporary healthcare system is being revolutionized by Deep Learning (DL) concepts, yet challenge for forecasting accurate cardiac illness and reliability. With medical data, it is hoped that combining the two procedures will improve the effectiveness of established strategies for predicting heart disease. The research that is being suggested presents a novel way for boosting heart illness forecasting preciseness while simultaneously integrating with nature-inspired optimization techniques that are associated with existing feature engineering methods. Hence, the combined approach of Binary Particle Swarm Optimization, assisted with Attention-based Deep Network (BPSO-ADN), has been suggested to efficiently extract significant features from the dataset to enhance prediction accuracy. The research uses a cardiac dataset followed by pre-processing, which needs to be considered. BPSO, used for feature selection, and ADN, used for analyzing detailed patterns within the data, provide a synergistic relationship at the heart of the technique. Throughout its repeated exploration of feature subsets, BPSO is directed by a fitness evaluation method to arrive at the subset most suitable for the prediction of heart disease. By utilizing self-attention processes, the ADN component can identify dependencies within the data, which ultimately results in higher predictive performance for both binary and multivariate classification of attributes.

Keywords: Binary Particle Swarm Optimization; Deep Neural Network; Heart Disease Prediction; Nature-Inspired Optimization; Self-Attention Model.

1. Introduction

In the world, heart diseases also known as cardio-vascular illness leads to illness and mortality, meaning that they are responsible for more than 70 percent of all deaths. According to the findings of the global burden of Disease study, heart disease makes up around 43 percent of all deaths [1]. The primary risk factors for cardiovascular disease include age, gender, tobacco use, lipids in the diet, lack of physical activity, high levels of cholesterol, being overweight, and consumption of alcohol [2]. Presently, research is being conducted to build technology with artificial intelligence [3] for evaluating cardiac illnesses and creating autonomous diagnosis methods for various other medical conditions. Machine learning (ML) and deep learning (DL)[4] approaches are transforming the existing healthcare system. However, effectively and reliably predicting cardiac illness remains a tough task.

The prompt detection of people who are at high risk for heart failure is one way that deep neural network models could help counteract the rising racial disparities in the outcomes of heart failure. However, not much research has been done on the demographic biases present in these models' efficiency [5]. AI-driven systems improve diagnostic accuracy in the field of medical imaging, which also provides assessments that are both quick and accurate [6]. Realistic, patient-specific simulations of normal and diseased hearts enable quantitative heart

function understanding and their use in various applications, such as improving evaluation, offering real-time assistance in making choices, forecasting future outcomes, and assisting with medical decisions [7].

However, the optimal selection of features from the dataset makes better decisions for heart disease prediction, leading to better accuracy in forecasting the disease as normal and abnormal using nature-inspired optimization techniques. Artificial life refers to replicating live behaviour observed in animals and people. It serves as a source of inspiration [8] for developing algorithms commonly employed in tasks like pattern recognition, classification, selection of features, optimization, and other related applications. Various fields have seen the emergence of swarm intelligence optimization algorithms as reliable approaches for feature selection. A form of AI known as swarm intelligence is based on collective behaviours in decentralized and self-organized systems [9]. The study on heart failure problems aims at enhancing heart failure forecasts with maximum accuracy, specifically by combining the deep neural network analysis using the nature inspired particle swarm optimization approach [10].

The most recent studies show the many approaches for forecasting the preciseness of heart illness. Researchers have used ensemble learning, feature extraction, and ML & DL models to improve prediction models' accuracy and effectiveness and have achieved amazing results [11]. Furthermore, the identification of features [12] is an essential component in the process of forecasting cardiac disease. When it comes to identifying potential patients through accurate prediction, the methods with a higher level of accuracy and reliability provide more assistance with the heart disease data. Compared with the prior algorithms with preliteracy of heart illness, the forecasting study uses a deep neural network with optimal sized technique effect.

The key contributions of this research include:

- To develop an integrated framework that includes BPSO for feature selection and ADN for accurate heart disease prediction, aiming to enhance predictive accuracy while using optimized feature engineering policies.
- To access the efficacy of the BPSO-AND model by conducting comprehensive experimentation on a heart disease dataset with binary and multiclass classification using various attributes.
- To contrast the current research study with other existing algorithms commonly used in heart disease prediction, using various performance metrics.

An outline of the research article is given below. In Section 2, the recent literature using various ML and DL algorithms is reviewed thoroughly. The implemented model strategy, research methodology and feature processing steps are detailed in Section 3. The results analysis with performance metrics is covered in the 4th Section. Section 5 explores the summary and scope for further enhancement.

2. Research Methodology

Bhatt et al. [14] developed a diagnosis model to accurately predict cardiovascular disease using clustering and other ML algorithms like decision trees, random forests, and XGBoost. GridSearchCV was utilized for enhancing employed variables to maximize outcome. The application of the method is used on a dataset of 70,000+ cases from Kaggle. Following an 80:20, for a split of the data for training, the model's accuracy reached 87.2%, and the precision was 0.95%, in accordance with study's findings. The result's interpretability and the algorithm's capacity to explain the clusters it created were not assessed.

Nagavelli et al. [15] proposed a quick analysis of heart disease detection ML models like naïve Bayes with a weighted method to predict cardiac disease. The second one automatically analyzes coronary heart condition detection using the frequency and time domain. This approach uses contrary methods. The third method uses an enhanced support vector machine based on its dual optimization technique to identify heart failure automatically. 38 PhysioNet ECG samples with heart disease tested the suggested approach and provided 95.9% accuracy, 97.1% precision, 94.6% recall, and 95.3% f1-measure.

The supervised learning model called the Learning Vector Quantization Technique (LVQT) for accurately predicting heart disease using the UCI dataset is analyzed by Srinivasan et al. [16]. Following the completion of the data collection procedure, data cleaning and pre-processing stages were followed. For predictive analysis, models based on machine learning are utilized after this phase. The system receives a set of training sequences with identifiable classifications and initial output variable allocation using an algorithm for learning. After training, LVQ organizes the vectors of inputs into the same category as output vectors. The results demonstrated that the study yields accurateness of 98%, preciseness of 0.97%, 97.9% sensitivity, and 97.8% score values, respectively.

Rahman et al. [17] deployed an innovative framework to forecast the risk of heart disease by combining Transformer Networks with Self-attention Mechanisms (TNSM). The layers responsible for self-attention gather contextual data and produce representations that effectively model intricate patterns within the data. By assigning a specific attention weight to each input sequence element, self-attention mechanisms enable interpretability. Gathering pertinent data entails altering the attention processes, adding new levels, and modulating both the input and the output layers, and the results with 96.5% accurateness.

Dileep et al. [18] suggested cluster-based bi-directional long-short term memory (C-Bi-LSTM) for accurately predicting heart disease. The datasets from the heart illness attributes were utilized to obtain result findings. The cluster models source to pre-process and remove duplicate data. Following this, the C-Bi-LSTM approach is used to predict heart illness. The system's accuracy, sensitivity, and F1 score are the methods helps for indicating the model outcomes. The findings suggest that the study is the most effective, achieving an accuracy of 94.7% along with 0.924%.

Mandava and Manjula [19] proposed a hybridized deep learning approach called improved DenseNet201, and a deep residual shrinkage model was employed to extract features and predict heart disease. By generating the most effective viable option while evaluation conditions, the resulting intelligent system for making feasible to clinicians to accurately identify heart illness victims with greater precision. The study results demonstrated using the UCI dataset and accurately predicting the disease with 99.1% accuracy.

3. Proposed methodology

The research intends to reduce the expenses of implementing complex engineering strategy to boosts result of classification findings. This is made possible by the provision of an efficient BPSO-ADN strategy. A cardiac dataset obtained from a freely accessible database serves as the focal point of the proposed research. Pre-processing of the data is performed before applying deep learning techniques. A self-attention deep neural network model can understand intricate patterns and correlations within the dataset, potentially leading to increased predictive performance compared to typical machine learning methods. When paired with other methods, such as BPSO, for selecting features, it forms a strong methodology that may be used to extract significant representations from complex data and make accurate predictions. The ability to accurately anticipate cardiovascular heart disease is driven by the fact that it has the ability for saving lives of individuals, enhance research findings, and more effectively disperse healthcare output.

a. Dataset and pre-processing

The proposed model is analyzed using the dataset acquired from [20], and the data source description is presented in Table 1.

Table 1: Data Source Description

Source Information	Description
Heart disease	UCI
Instances	303
Type	Multivariate
Formatting file	.xls

The information presented in Table 2 pertains to the prediction of heart disease using 14 attributes stored in the Cleveland database. These attributes are arranged in various categories, ranging from 0 indicates normal condition and 1, 2, 3, and 4 indicating the presence of disease.

Table 2: Heart disease data attributes

S. No	Attributes	Description / Unit
1	Aging	yrs
2	Gender	1: male, 0: female
3	Chest illness	1- typical angina, 2 -atypical angina

4	blood pressure	3 -non-angina, 4- asymptomatic
5	Cholestral	mm/hg
6	Fasting	mg/dl
7	electrocardiograph	fbs>120
8	oldpeak	1: abnormal, 0: Normal
9	slope	stress
10	ca	1: upslope, 2: flat, 3: downslope
11	thal	Major count (0-3) fluoroscopy
12	num	3: normal, 6: fixed, 7: reversible
		Disease prediction status 0:<50%,1: >50%

A significant number of the values in the dataset are absent; therefore, to ensure that the dataset is considered dependable and of excellent quality, it is necessary to locate the missing values. If the absence of information is not handled appropriately, it has the potential to yield skewed results and conclusions that are misguided. Determine the mean of the subsequent data points for each attribute lacking values. The process involves adding up all of the now available numbers and then restricting the results based on the amount of data that is not missing.

The dataset has been pre-processed for encoding category variables, handling missing values, and normalizing numerical characteristics. Examining each attribute can help you locate any values that are missing. The Z-scoring z_a gives the variance throughout an medical condition, which estimates the dissimilarity between x_a and μ_a . With the assistance of this strategy, the model can correctly identify and handle extreme values that appear in the data.

$$z_a = \frac{x_a - \mu_a}{\sigma_a} \tag{1}$$

Where in equation (1), z_a refers to the standardized value of the ath attribute, x_a denotes the original value of the ath attribute, μ_a gives mean of the ath attribute and σ_a refers to the standard representation of the ath attribute.

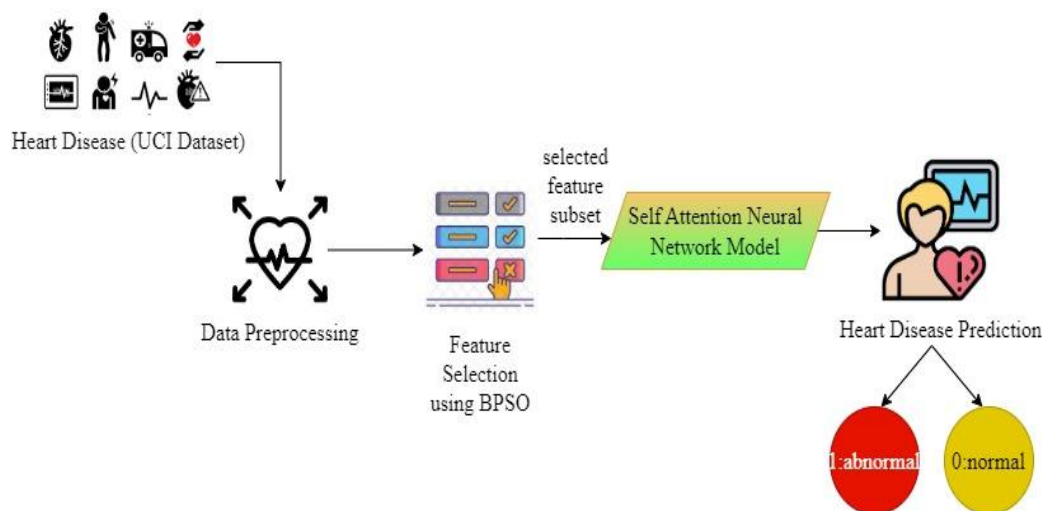


Figure 1: Overall Proposed Architecture

Figure 1 shows that the layers of EBDPF Architecture are responsible for managing the synchronization of data collected from various sources. Subsequently, it ensures the data is integrated and updated seamlessly for analysis and processing. Integrating with a wide variety of business systems enables the EBDPF to simplify the flow and processing of data smoothly. Enhancing the user experience and making system maintenance easier are two benefits that may be achieved through visualization tools and user-friendly interfaces. The EBDPF prioritizes scalability and performance optimization, which enables the platform to manage large-scale data operations and adapt to different data volumes and processing nodes. Therefore, this guarantees that the framework can process and analyze data in educational contexts efficiently and reliably while retaining high efficiency.

b. Feature Selection Using BPSO

The initialization of particles involves a random binary representation of each particle followed by $fitness(x_a)$ where x_a defines the binary representation of the selected features for the a th particle. L1 regularization is a technique that improves the effectiveness of the logistic regression model, which is well-known for its effectiveness in binary classification issues. A penalty term that is comparable to the true measurement of the magnitude of the coefficients is introduced by this regularization procedure. L1 regularization tends to create sparse solutions, which inherently perform computation with coefficients of less significant to zero. This is the fundamental advantage of adding L1 regularization. A swarm of agents, also known as particles, travels over the search area to find the best possible answer. Speed, position, and fitness value are the three parameters associated with each particle. Each atom and particle can preserve a record of its finest answer, which is the best in its class. First, let's make some assumptions regarding the factors being considered. Create a "population" of agents (particles) distributed evenly throughout the population. Determine the location of each particle with the fitness function through analysis. When a particle's current position is advantageous compared to its previous best position, it is deserving of being updated. Find out which particle is the best possible option. Improve the particle velocities by using the equation technique. Figure 2 explains the flowchart of BPSO.

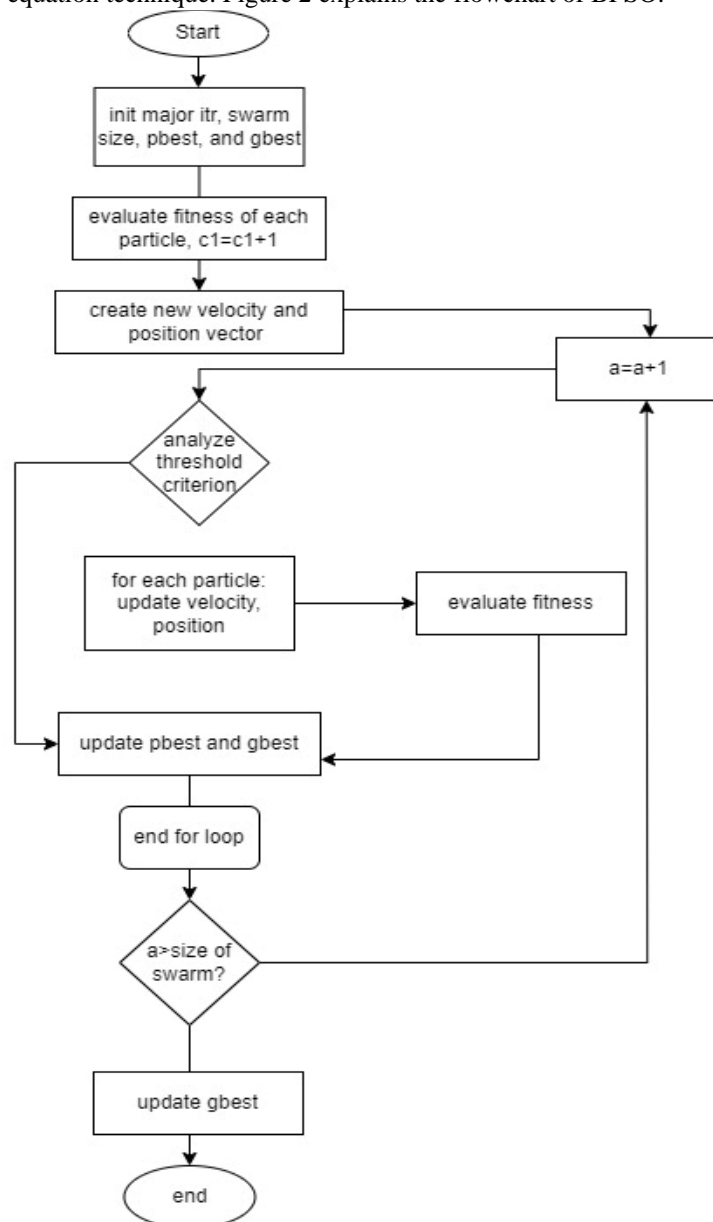


Figure 2: Flowchart of BPSO

velocity update equation is given as

$$v_{ab} = v_{ab} \cdot \text{sig}(v_{ab} + k_1 \cdot \mathfrak{R}_1 \cdot (pbest_{ab} - x_{ab}) + k_2 \cdot \mathfrak{R}_2 \cdot (gbest_b - x_{ab})) \quad (2)$$

$$x_{ab} = \text{sig}(v_{ab}) \quad (3)$$

From equations (2) and (3), v_{ab} represents the velocity of the ath particle for the bth feature and indicates how much the particle's position will change in the next iteration. x_{ab} represents the current position of the ath particle for the bth feature and describes the binary form of feature selection during the current iteration. The input values are compressed to fall within the range of 0 to 1 by the sigmoid activation function shown above. $pbest_{ab}$ refers to the personal best position of the ath particle for the bth feature, representing the best feature subset that the particle has found so far during the optimization process. $gbest_b$ defines the global best position for the bth feature among all particles in the swarm and feature subset in the swarm. k_1 and k_2 represents the acceleration coefficient that controls the impact of $pbest_{ab}$ and $gbest_b$ positions on the particle's velocity update. They determined to strike a balance between exploration, which involves discovering new solutions, and exploitation, which consists of using proven effective solutions. From 0 to 1, these are values that are chosen at random, and it is represented using \mathfrak{R}_1 and \mathfrak{R}_2 . Incorporating stochasticity into the equation for updating the velocity makes it possible for the particles to do more efficient explorations of the search space. The purpose of its application in this context is to guarantee that the velocity values are kept within the acceptable range of 0 to 1, which is appropriate for the binary form in BPSO.

Through feature selection, BPSO improves heart disease prediction accuracy. BPSO's iterative nature explores heart disease feature subsets. In every iteration, feature subsets are assessed for their heart disease prediction accuracy. BPSO refines the search space by updating the motion and location of particles representing distinct feature subsets concurrent to the optimal subset. Fitness evaluations only select the most informative information, improving performance prediction. BPSO's global best position is the feature subset that optimizes heart disease prediction. So, BPSO is an effective optimization technique that helps heart disease models of prediction locate the most important variables for accurate diagnosis.

Based on the fact that it directs the search process for an optimal selection of features from the dataset, the velocity update equation in BPSO is connected to the heart disease prediction. Regarding heart disease prediction, each feature subset is represented as a binary string, with each bit indicating whether a feature is included or excluded from the set. The feature subsets acquired by BPSO are utilized to train a model for predicting heart disease, such as a classifier. The set features contribute to developing an algorithm that can identify individuals at risk of illness related to heart with selected features.

c. Model Architecture

In the ADN model the input is necessary to represent the selected features. When preparing the feature sequences to be input into the Self-Attention model, it is recommended to use techniques such as source embed and positional selection. The ADN mechanism should be utilized to identify the connections between the different components of the input component sequences. It is necessary to use numerous parallel Self-Attention layers to capture the many dependencies and interactions inside the feature sequences. A deep neural network with feed-forward algorithms should be used to process the outputs of the Individual attention layers to do non-linear processing and make the final prediction.

During this research, the model is applied for training that involves multiclass and binary classification objectives. To solve the challenge of binary classification, the result provides predictions regarding whether there is a normal (0) or abnormal (1) category of "restecg" attribute of cardiac disease. To solve the challenge of four-class classification, the partition of labels into four distinct categories: that correspond to the different levels of risk for cardiovascular disease. The input representation of features denoted as X in equation (5) shows features in the sequence.

$$X = [x_i, x_{i1}, \dots, x_{in}] \quad (4)$$

Where in represents the length of the input sequence, x_a represents the feature in the sequence. The input embedding function in equation (5) is defined as x'_a :

$$x'_a = \text{Concat}(e(x_a), f(x_a)) \quad (5)$$

Where $e(x_a)$ represents the embedding of categorical variables and $f(x_a)$ represents the scaled numerical features.

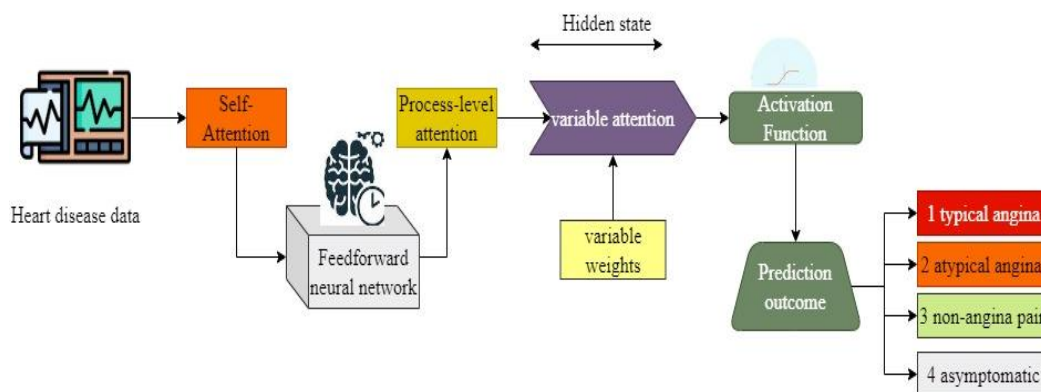


Figure 3: Working of Heart Illness Forecast Model

Utilizing data on heart disease, the Self-Attention deep neural network model is combined with the chosen attributes, as shown in Figure 3. The performance should be evaluated with a different test dataset, with measures to predict heart disease. The result integrated technique is compared with various feature selection methods and models.

d. Performance Evaluation

Following the completion of the model architecture for the prediction of heart illness that uses the combined BPSO-ADN, the subsequent step of the utmost importance is carefully evaluating its performance. Several technical procedures are utilized in this phase to evaluate the effectiveness and resilience of the constructed model. The model uses the heart dataset to carry out A comprehensive training regimen. During this training, the model acquires knowledge of complicated patterns and correlations between the input features and the illness status. By doing continuous training epochs, the model's variables and hyperparameter settings are fine-tuned to improve predicted accuracy and optimize performance. In model training, the number of training epochs controls the dataset iterations. At the same time, optimizers and loss functions are responsible for guiding the parameter updates. It is essential to fine-tune these parameters to optimize the model's performance. It is necessary to use approaches such as random search to determine the ideal values tailored to the dataset and the available computational resources. The number of model training epochs as particles utilized in BPSO controls the no. of dataset epochs in model. At the same time, optimizers and loss functions are responsible for guiding the parameter updates.

4. Results and Discussion

For implementing the research, Python Language is used for its extensive library TensorFlow and PyTorch for modern ML and DL model and optimization libraries. Depending on the dataset's amount and the model's complexity, this comprises an appropriate graphics processing unit (GPU) for accelerated computations. The existing algorithms, such as LVQT [13], TNSM [13], and C-Bi-LSTM [14], are taken for comparison study with the following metrics: accuracy, precision, recall and f1-score analysis for enhancing heart disease accuracy.

a. Accuracy

A metric called accuracy predictions, accuracy is used in heart disease prediction. A high accuracy suggests that the model accurately categorizes the majority of cases into the categories that are appropriate for them, for example, whether or not heart disease is present, either normal or abnormal, in binary classification. The researcher's objective is to achieve the highest possible level of accuracy to guarantee that the model can accurately identify individuals at risk of developing heart disease.

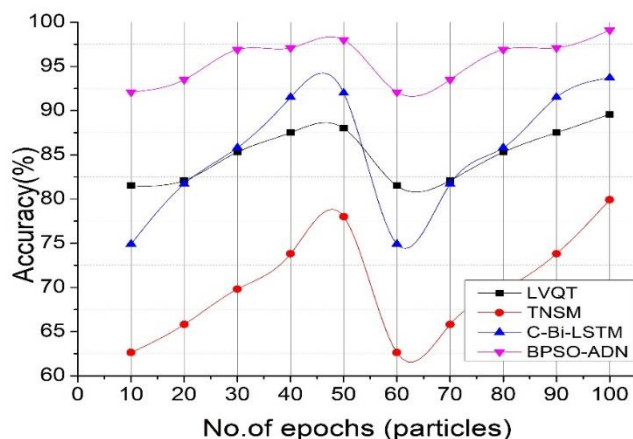


Figure 4: Accuracy Analysis

In Figure 4, the high accuracy of the BPSO-ADN technique suggests that the model can categorize individuals reliably as either having heart disease or having normal health behaviour without heart disease according to the features that have been collected. It is an important metric of the model's efficacy in clinical practice and represents its capacity to produce accurate predictions. Accuracy is a measurement that expresses how accurate the projections the model makes are. It is the percentage of examples correctly classified out of the overall number of instances.

b. Precision

In predicting heart disease, precision refers to the capacity of the model to accurately identify individuals with cardiovascular disease among those who are anticipated to have a positive result. The model has a low rate of incorrect diagnoses, meaning it has a high precision. This reduces the likelihood that individuals would be wrongly diagnosed with heart disease when, in fact, they do not have the condition.

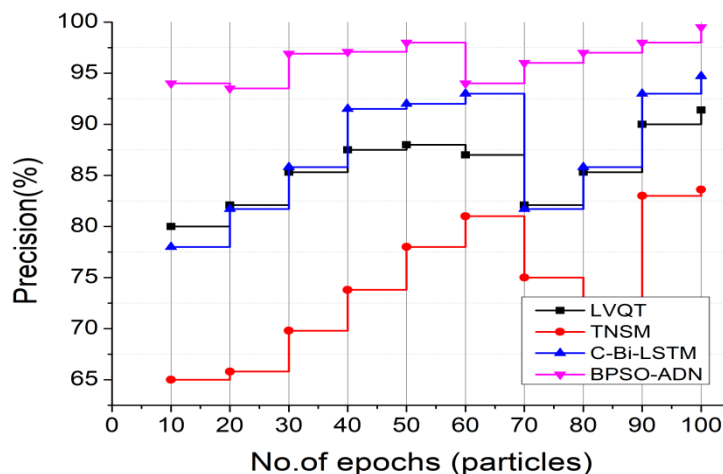


Figure 5: Precision Analysis

On the other hand, the precision shown in Figure 5 is concerned with the accuracy of the positive predictions that the model has produced. When researching the forecasting of heart illness, limiting the number of false positives is of utmost importance. This is because incorrectly labelling healthy patients as having the cardiovascular disease can result in unneeded stress and the need for medical measures.

c. Recall

The recall metric evaluates the capacity of the model to accurately identify individuals diagnosed with cardiovascular disease out of the whole population who have the ailment. If the model has a high recall, it can catch the majority of cases of heart disease successfully, hence reducing the likelihood of false negatives, which are cases that are overlooked.

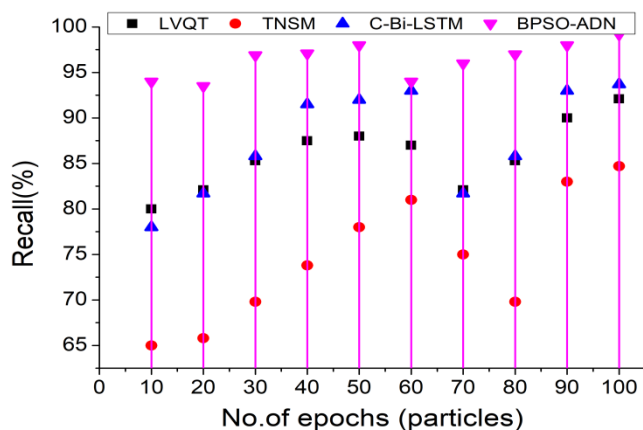


Figure 6: Recall Analysis

In figure 6, the graph shows to catch all true events, including actual condition of heart illness, is measured by this criteria. When researching the forecasting of heart illness, reducing the number of false negatives is of the utmost importance. This is because the failure to identify individuals who have heart disease can lead to delayed treatment and severe health effects.

d. F1-Score Analysis

As a result of considering both erroneous true and error as negatives, the mentioned metric finds a happy medium between two metrics, as shown in Figure 7. A high F1 score indicates that the model reaches both high accuracy and high recall, which makes it an accurate indicator of heart disease. This is the case when predicting risk factors for heart disease. To develop reliable models for predicting cardiac disease that are capable of achieving both high accuracy and high recall, researchers are working to maximize the F1 score. A high F1 score indicates that the model can accurately identify individuals at risk of developing heart illness while decreasing the number of erroneous predictions. This ensures that the model is reliable when applied in clinical settings.

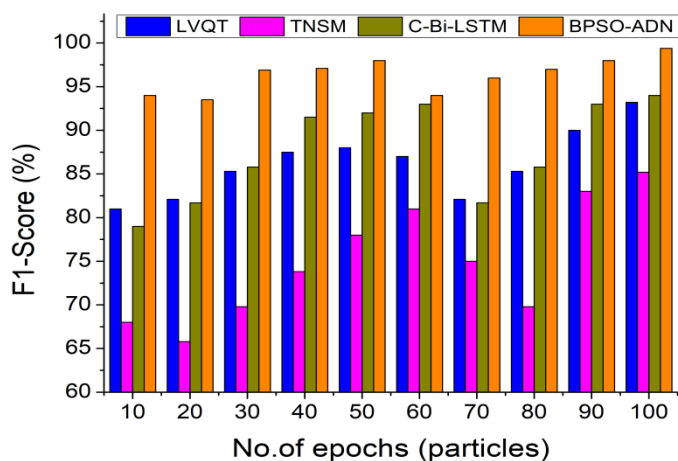


Figure 7: F1-Score Analysis

The evaluation of the BPSO-ADN method for heart disease prediction. High accuracy means the model can correctly categorize people as either suffering from or not suffering from heart disease as 0 for normal and 1 as abnormal. At the same time, accurate information shows that positive cases are at risk, limiting false positives. High recall guarantees the program captures most heart disease patients, reducing missed diagnoses. The result indicates the model's ability to identify heart disease risk factors with few erroneous predictions reliably. They provide a complete assessment of the BPSO-ADN approach's heart disease prediction ability, improving patient results and healthcare decision-making.

5. Conclusion

The proposed research presents a novel approach to the prediction of cardiac illness. This strategy is achieved by combining BPSO-ADN. This combination framework offers a promising path for significantly lowering the expenses associated with feature engineering while simultaneously improving classification accuracy. Through the utilization of BPSO for selecting features and ADN to collect detailed patterns throughout the training set, the algorithm displays the ability to surpass conventional machine learning techniques. Comprehensive experimentation and evaluation have been conducted to demonstrate the efficacy of the BPSO-ADN technique, which reflects its capacity for forecasting the heart illness properly. By anticipating heart sickness based on clinical information and the patient's severity level, DL algorithms can potentially lower the number of fatalities. During the optimization process, suitable feature variables are selected to produce the most accurate prediction possible for a sufficient sample size of generated samples. To address difficult challenges in the healthcare industry, the study highlights the significance of utilizing innovative optimization techniques and developing DL architectures. This will ultimately lead to enhanced outcomes for patients and resource allocation in medicine.

The future scope of this research focuses on during the process of optimization, developing adaptive mechanisms that can dynamically modify BPSO parameters such as the inertial weight and accelerator coefficients in future to enhance the efficiency of the algorithm. The utilization of dynamic parameter adaptation with convergence and boosting the algorithm's adaptability to various problem environments. Incorporating long-term data and continuous monitoring into the research could make it easier to construct predictive models that are capable of early diagnosis and individualized intervention, which would ultimately lead to advancements in the field of heart condition monitoring and management.

References

- [1] Drożdż, Karolina, et al. "Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: a machine learning approach." *Cardiovascular Diabetology* 21.1 (2022): 240.
- [2] Juhola, Martti, et al. "Data analytics for cardiac diseases." *Computers in Biology and Medicine* 142 (2022): 105218.
- [3] Saqlain, Syed Muhammad, et al. "Fisher score and Matthews correlation coefficient-based feature subset selection for heart disease diagnosis using support vector machines." *Knowledge and Information Systems* 58 (2019): 139-167.
- [4] Nouman, Aleeza, and Salman Muneer. "A systematic literature review on heart disease prediction using blockchain and machine learning techniques." *International Journal of Computational and Innovative Sciences* 1.4 (2022): 1-6.
- [5] Kaur, Dhamaanpreet, et al. "Race, sex, and age disparities in the performance of ECG deep learning models predicting heart failure." *Circulation: Heart Failure* 17.1 (2024): e010879.
- [6] Rana, Md Shohel, and Jeff Shuford. "AI in Healthcare: Transforming Patient Care through Predictive Analytics and Decision Support Systems." *Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006-4023* 1.1 (2024).
- [7] Cicci, Ludovica, et al. "Efficient approximation of cardiac mechanics through reduced-order modeling with deep learning-based operator approximation." *International Journal for Numerical Methods in Biomedical Engineering* 40.1 (2024): e3783.
- [8] Nugroho, Andi, et al. "Comparison of binary particle swarm optimization and binary dragonfly algorithm for choosing the feature selection." *2021 5th International Conference on Informatics and Computational Sciences (ICICoS)*. IEEE, 2021.
- [9] Rostami, Mehrdad, et al. "Review of swarm intelligence-based feature selection methods." *Engineering Applications of Artificial Intelligence* 100 (2021): 104210.
- [10] Yuliandari, Dewi, et al. "Improving the Accuracy of Heart Failure Prediction Using the Particle Swarm Optimization Method." *Sinkron: jurnal dan penelitian teknik informatika* 9.1 (2024): 210-220.
- [11] Khan, Arsalan, et al. "A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction." *Health & Social Care in the Community* 2023 (2023).
- [12] Budholiya, Kartik, Shailendra Kumar Shrivastava, and Vivek Sharma. "An optimized XGBoost based diagnostic system for effective heart disease prediction." *Journal of King Saud University-Computer and Information Sciences* 34.7 (2022): 4514-4523.
- [13] Hassan, Diman, Haval I. Hussein, and Masoud M. Hassan. "Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis." *Biomedical Signal Processing and Control* 79 (2023): 104019.
- [14] Bhatt, Chintan M., et al. "Effective heart disease prediction using machine learning techniques." *Algorithms* 16.2 (2023): 88.
- [15] Nagavelli, Umarani, Debabrata Samanta, and Partha Chakraborty. "Machine learning technology-based heart disease detection models." *Journal of Healthcare Engineering* 2022 (2022).
- [16] Srinivasan, Saravanan, et al. "An active learning machine technique based prediction of cardiovascular heart disease from UCI-repository database." *Scientific Reports* 13.1 (2023): 13588.
- [17] Rahman, Atta Ur, et al. "Enhancing heart disease prediction using a self-attention-based transformer model." *Scientific Reports* 14.1 (2024): 514.

-
- [18] Dileep, P., et al. "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm." *Neural Computing and Applications* 35.10 (2023): 7253-7266.
- [19] Mandava, Manjula. "MDensNet201-IDRSRNet: Efficient cardiovascular disease prediction system using hybrid deep learning." *Biomedical Signal Processing and Control* 93 (2024): 106147.
- [20] <https://archive.ics.uci.edu/dataset/45/heart+disease>