International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V73I3P135

Volume 73 Issue 3, 503-516, March 2025

Original Article

Heart Disease Prediction Using a Novel Fuzzy-Enhanced CLSTM Model with Adaptive Stochastic Gradient Descent Optimization

R. Parthiban[1], K. Santhosh Kumar[2]

[1]*Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.*
[2]*Department of Information Technology, Annamalai University, Chidambaram, Tamilnadu, India.*

[1]*Corresponding Author : parthineyveli@gmail.com*

Received: 22 October 2024 Revised: 15 February 2025 Accepted: 20 February 2025 Published: 28 March 2025

Abstract *- Cardiovascular Diseases (CVDs) are considered to be the predominant cause of the increase in death rates around the world. Hence, early detection is mandatory for managing and providing the affected persons with effective treatment. Normally, capturing complex patterns in medical data is much more difficult with the help of traditional machine learning methods. Although it is more effective, it is not able to handle the uncertainty and non-linearity problem that exists in the patient's health metrics. In this study, a novel approach called Fuzzy-Enhanced CLSTM is proposed for heart disease prediction. This novel approach integrates Multilayer Fuzzy-based Convolutional Neural Networks (MFCNN) in correlation with the Bidirectional Long Short-Term Memory (BiLSTM) model. In this method, the fuzzy logic is leveraged to enhance the feature extraction process of CNN by making it more robust in dealing with imprecise and uncertain data. Including fuzzification will enhance the sensitivity to a certain extent by supporting the critical variations in the clinical parameters. Combining convolutional neural networks with the BiLSTM will capture the temporal dependencies in sequential data, enabling a more comprehensive understanding of patient history and trends over time. Hence, this model is suitable for both spatial feature extraction and temporal analysis. Fine-tuning of the model is performed using an Adaptive Stochastic Gradient Descent (ASGD) optimizer, which dynamically adjusts the learning rate during training. This helps faster convergence and prevents the model from getting stuck in local minima by improving overall prediction accuracy. The experimental results conducted by using publicly available datasets provide significant improvement in early and accurate heart disease detection and prediction by providing better accuracy and generalization compared to other traditional Methods.*

Keywords *- Heart disease prediction, Cardiovascular diseases, Fuzzy-based CNN, Adaptive stochastic gradient descent, BiLSTM, Deep learning.*

1. Introduction

Cardiovascular disease remains a prominent global health issue, resulting in a substantial number of deaths each year. Having access to various large datasets about healthcare and fine-tuning the methods of machine learning that are more recent and sophisticated offer the opportunity to enhance the precision of cardiac disease diagnoses [1]. Over the course of 2019, the World Health Organisation (WHO) reported that 18.6 million people around the world passed away as a direct result of heart disease. The number of lives lost due to cardiovascular disease is greater than that of any other single cause worldwide [2]. There is a need to detect the patient's condition in the early stages, and to reduce mortality rates, an accurate prognosis is needed to detect heart disease at the earlier stages. Considering these factors, there is insufficient information about the patient's risk conditions, and it is mandatory to check with appropriate methodologies. Machine learning is used to analyze complicated medical data taken in large quantities. This will help to identify the patterns that the traditional methods can't do [3]. Some examples of the types of datasets that can be analyzed by these methods include electrocardiograms (ECGs), medical images, patient histories, and laboratory results. These datasets contain a lot of information that assists the healthcare department professionals in researching the specified area using machine learning models and increasing the accuracy of predictions. From machine learning models, the supervised learning algorithm is the only approach commonly used for predicting cardiovascular disease [4]. The dynamic and complex patterns that can be found in medical data are beyond the capabilities of supervised learning models, even though these models show promising methods that place a primary emphasis on

sequential decision-making as the primary methodology. The optimization of a cumulative reward signal is accomplished by reinforcement learning agents through a series of decisions [5]. This model has been successfully implemented in various fields, including gaming, robotics, and recommendation systems. In the field of medicine, reinforcement learning is supported in the analysis and prediction of various forms of the disease since it has the maximum potential to enhance the prediction efficiency in various protocols used for treating CVD patients with the support of new medical interventions [6]. The improvement is made possible with the possible prediction phenomenon by using reinforcement learning with better strategy by taking the sensitive details of the patients by analyzing their responses in correlation with the useful decisions.

This could be possible by the implication of the hybrid methodology for both prediction and classification. Nowadays, data augmentation is a widely used methodology for CVD prediction with the help of natural language processing and computer vision technologies [7]. When providing patients with individualized healthcare, reinforcement learning models are an excellent option to consider. This is because of its decision-making effectiveness and adaptability capabilities. The diversity of the training data is being increased to make the models more accurate. Normally, data augmentation is used to enhance the data related to healthcare by creating fictitious data points that look like the actual patient data. The problem with the imbalanced datasets, scarcity, and privacy-related issues are resolved by using this approach [8]. By incorporating artificial data into the training dataset, machine learning models can enhance their capacity to identify patterns and characteristics.

The problem in the accurate prediction of cardiovascular disease significantly impacts the health of the general population and the quality of care provided to patients. This happens because of the increase in the number of cardiovascular disease-affected patients worldwide, which is considered to be a major concern in terms of global health. Comparing the large datasets obtained from the healthcare industry with state-of-the-art machine learning techniques may improve the prediction of cardiovascular diseases [9]. Conventional risk assessment techniques typically consider a patient's age, gender, and lifestyle choices, among other demographic data; nevertheless, it is possible that these factors are not at all relevant. Therefore, it is crucial to identify the patterns disregarded by conventional methods to obtain a dependable prognosis. This is because medical data is extremely complex, and dealing with such data is difficult. Machine learning [10] makes it possible to analyze highly complex medical data and recognize patterns essential for accurate disease prediction. Diseases of the cardiovascular system, arrhythmias, and heart failure are all examples of conditions that should be considered under this umbrella. It is difficult to make reliable predictions about cardiovascular disease because of the gathering of variables that can influence the condition [11]. The flexibility of reinforcement learning algorithms is an important consideration in our expedition for greater precision. These algorithms must improve and modify their prediction models for fluctuating patient data [12]. These innovative methods combine reinforcement learning with normal data augmentation techniques by increasing the accuracy of predictions regarding heart disease.

2. Related Works

In recent years, there has been a notable rise in machine learning and artificial intelligence use to predict medical issues. Various analyses are made in medical data to determine how well these methods work in bringing attention to diagnosing various heart diseases accurately and faster. This is achieved by comparing the different algorithms and their effectiveness in detecting potentially dangerous heart conditions. The work [13] examines several different models concerning their ability to forecast the probability of occurrence of cardiac problems. To improve the accuracy of the predictive models designed for cardiac problems, it is necessary to use different machine-learning approaches.

The study [14] explores possible ways to evaluate the efficiency of detection and prediction by using these algorithms to manage medical data and facilitate accurate diagnosis. Using data mining techniques to recognize significant patterns in patient records is important, which could result in improved early warning signs of cardiovascular problems. By employing strategies driven by data, the research brings about improvements in patient care and sheds light on predictive modeling for hospital readmissions [15]. In order to make the most of a limited dataset, it is necessary to supplement it with every conceivable combination of input features.

Through the use of the Mix-up augmentation method, it was suggested that the model's capacity to generalize could be improved. The creation of synthetic training examples from pairs of input samples and their labels is accomplished through linear interpolation in this method [16]. Through the process of combining image patches, the Cutout method was enhanced, which in turn promoted the learning of models from composite samples, resulting in improvements in both localization and classification. The Auto Augment algorithm was developed to reduce the amount of required human intervention. It identifies effective data augmentation policies on its own through autonomous means. The study [17] provides an investigation into the various methods of data enhancement for super-resolution images was carried out, and a novel approach to graphic enhancement was suggested. The various analyses of the literature are shown in Table 1.

Table 1. Review of literature

Author	Methodologies Used	Advantages	Limitations
Saranya et al. [18]	ML classification, feature optimization with sensitivity analysis.	Enhanced feature selection and improved prediction accuracy (98.45%).	Limited testing on diverse datasets.
Revathi et al. [19]	Optimized LSTM using Salp Swarm Algorithm and Genetic Algorithm.	High accuracy (97.11%), efficient feature elimination.	It may require extensive computational resources.
Darolia et al. [20]	LSTM combined with Quantum Neural Networks (QNN).	High prediction accuracy due to dimensionality reduction techniques and robustness.	Complex model structure requiring advanced tuning for specific datasets.
Oyewola et al. [21]	Ensemble deep learning using the Kaggle dataset.	High accuracy (98.45%), improved interpretability, and real-time diagnostic capabilities.	Dependence on pre-existing data, challenges with real-world data generalization.
Ramkumar et al. [22]	CNN-RNN architecture using IoMT data (e.g., pulse oximeters, ECG).	Real-time monitoring and high precision in continuous patient monitoring.	Limited scope in healthcare settings lacking IoT infrastructure.
Akhtar et al. [23]	Ensemble DL models, feature fusion.	Superior prediction accuracy through combining multiple models.	Computationally intensive and risk of overfitting.

The study [24] explains how Natural Language Processing (NLP) models can be made more resilient and generalizable through the use of data augmentation. The integration of multiple data augmentations is for improving the robustness of models and the assessment of uncertainty. These methods give us information about how to make more reliable and useful models in a wide range of situations. This study [25] describes recent developments, a discussion of obstacles, and opportunities for deep learning approaches to identify imbalanced data by examining the data augmentation techniques that are tailored to the detection of small objects, particularly when dealing with circumstances in which the objects being detected are of a small size. This study [26] provides various data augmentation techniques that utilize deep learning algorithms for the purpose of audio classification.

This study [27] investigates various techniques for improving data developed specifically to train recurrent neural networks for voice recognition devices. In order to improve the classification of endoscopic images, this work investigates data augmentation techniques. It provides a summary of the methodologies that are currently in use, as well as the impacts that they have. In the study [28], various data augmentation techniques are investigated within the context of autonomous vehicles. The study [29] also investigates how these techniques can be utilized to enhance the precision of sensor data that is utilized by autonomous driving systems. In order to improve investment strategies and address challenges in the financial sector, the author [30] presented a framework for deep reinforcement learning that uses RL methodologies. One of the most important aspects of the framework is the management of financial portfolios. The study demonstrates the possibility of deep reinforcement learning in the field of personalized medicine by demonstrating its ability to improve treatment decisions through the utilization of data obtained from medical registries. The dynamic treatment regimens are the primary focus of the research being conducted so far. The study [31] provides a deep Q-learning method that improves autonomous agents' decision-making capabilities by utilizing experiential learning. The method focuses on autonomous systems. The deep reinforcement learning algorithm [32] also provides possible ways to support the application of robotic radiation adaptation to treat lung cancer. This investigation is driven by the desire to improve patient outcomes by refining treatment regimens, which is the driving force behind this investigation. Both the benefits and drawbacks of implementing Deep Reinforcement Learning (DRL) in healthcare settings are discussed in this study.

The study [33] proposed a framework that employs deep reinforcement learning to achieve two objectives, namely, improving hemodynamic interventions and patient outcomes in critical care settings. According to the study findings, patients experiencing septic shock may be able to find individualized and flexible treatment options with DRL. The research in [34] examines the ways in which the staff of intensive care units is utilizing reinforcement learning algorithms to develop individualized treatment plans for patients who are critically ill and suffering from sepsis. The authors present a model that makes use of reinforcement learning to develop individualized treatment plans for patients. The purpose of this model is to improve the efficiency of interventions that are carried out in the intensive care unit. The study [35] provides a model that considers patient feedback for ventilation setting adjustments in order to accomplish this goal. Consider the example mentioned in large-scale medical research projects, which started with the aim of making the prediction more accurate, precise, and unique. The programs implemented in intensive care units might include both diagnostics and interventions.

This study demonstrates how reinforcement learning can improve mechanical ventilation techniques. The objective of this study is to improve the prediction outcome of the patient.

The data augmentation and reinforcement learning algorithm is used to solve the problems created in understanding cardiac health data. Traditional machine learning algorithms are unsuitable for predicting the complex patterns and non-linear correlations associated with medical data [36]. Data augmentation aims to improve the dataset by adding more examples from a larger variety of classifications. Models will be able to produce representations that are more thorough as a result. Applying fuzzy logic and reinforcement learning leads to enhancements in the capacity for sequential decision-making and easier adaptation to the dynamic field of heart health development [37-39]. The main objective of this integration, which seeks to enhance patient outcomes, is to increase prediction accuracy.

3. Proposed Methodology

The proposed method integrates an FCNN, which is used for feature extraction, a BiLSTM network, which is used for sequential data analysis, and an Adaptive Stochastic Gradient Descent (ASGD) for fine-tuning. The architecture of the proposed CLSTM approach with ASGD is illustrated in Figure 1. The input consists of patient health records, which include various clinical parameters (e.g., age, cholesterol level, blood pressure, etc.) taken from the dataset.

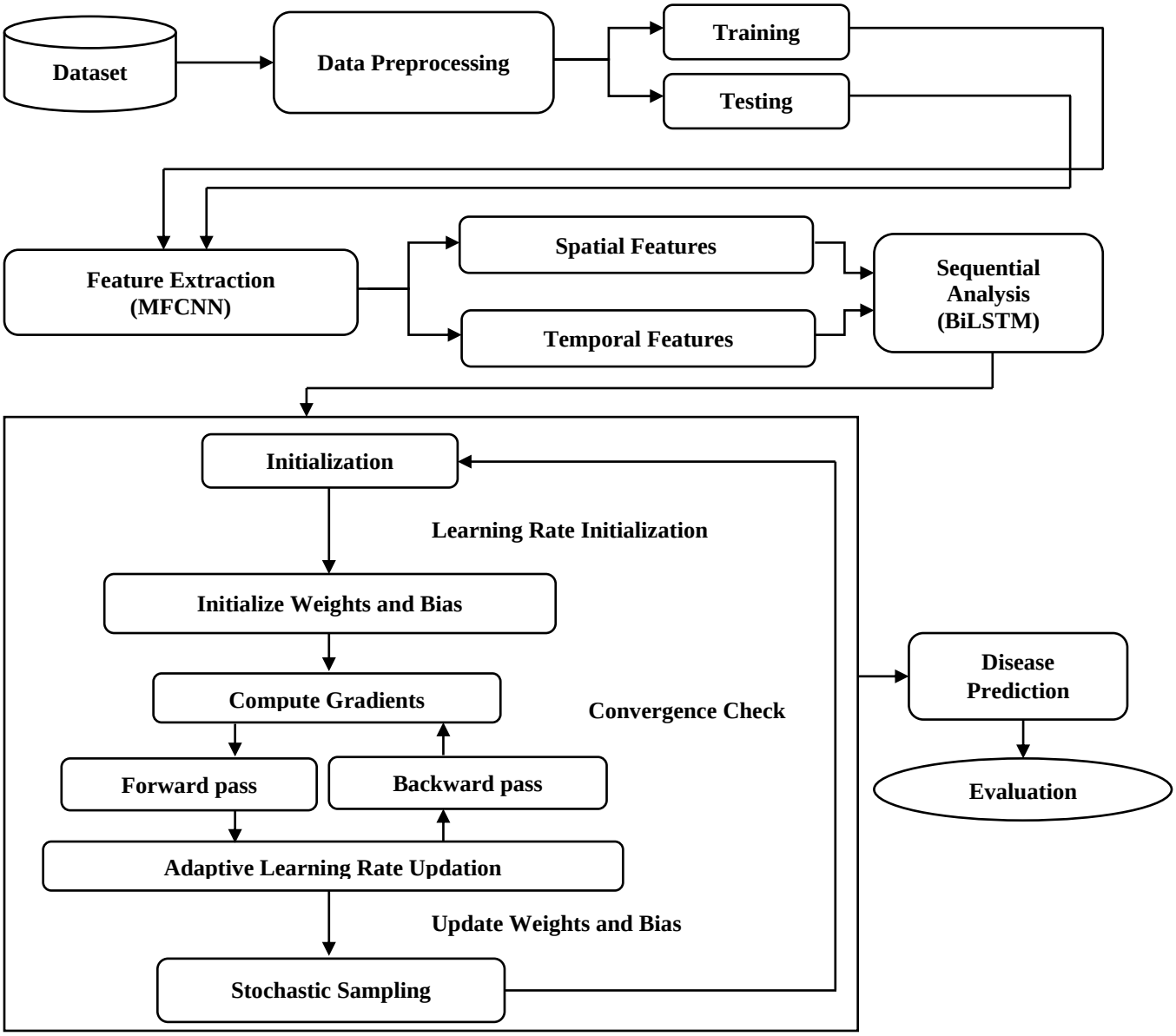

Fig. 1 Architecture of the CLSTM-ASGD system

3.1. Dataset Description

The total dataset taken for the analysis is 90,000 (patient data), which includes 11 different features. The dataset is taken from https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset/ data. In total, three various types of input features are taken for analysis. Initially, factual information is taken, then the information from the medical examination results is considered, and finally, the patient's related information is collected manually from direct analysis. The features taken are shown in Table 2. All the data have objective features, examination features, and subjective features. Age, height, and gender are objective features, cholesterol glucose are examination features, and smoke alcohol are subjective features.

The clinical parameters (features) and their considered ranges are explained in Table 2. The input data is then subjected to preprocessing, ensuring the raw medical data is clean, transformed, and prepared to extract meaningful patterns. The input data are taken and are considered as $I=\{a_1,a_2,...,a_n\}$where ai represents the ith feature of the patient record. The input layer also contains a fuzzification layer, which transforms the crisp input values into fuzzy values to handle uncertainties and improve feature representation. Both a training dataset as well as a testing dataset have been created from the data. The output from the data splitter is then passed to the MFCNN for further processing.

Table 2. Dataset details

S.No	Clinical Parameters	Feature Type	Range of Values
1	Age	Objective	18-60 years
2	Height	Objective	140-178 CMs
3	Weight	Objective	50-120 Kgs
4	Gender	Objective	Male: 0 & Female: 1
5	Systolic blood pressure	Examination	90-160
6	Diastolic blood pressure	Examination	60-100
7	Cholesterol	Examination	1-Normal; 2-Above normal; 3-Good
8	Glucose	Examination	1-Normal; 2-Above normal; 3-Good
9	Smoking	Subjective	Yes-1 & No-0
10	Alcohol consumption	Subjective	Yes-1 & No-0
11	Physical activity	Subjective	Yes-1 & No-0

3.2. Feature Extraction Using MFCNN

The fuzzy logic is incorporated into the CNN to handle the uncertainties in the obtained medical data. Here, for each instance of input feature ai, fuzzification is applied to transform the crisp values into fuzzy values. The fuzzy membership degree function is represented as, Membership Function, $M=m_A(a_i)+m_B(a_i)$

$$M=\frac{1}{1+\left(\frac{a_i-\eta_A}{w_A}\right)^2}+\frac{1}{1+\left(\frac{a_i-\eta_B}{w_B}\right)^2} \quad (1)$$

Where M is considered the membership function and (mA, mB) is represented as the membership degree of ai when fuzzy sets A and B are considered, one for spatial features and another for temporal features. The center of the fuzzy set is marked to be η, and the two sides of the center are represented with (wA, wB), which mentions the corresponding width of the dataset. CNN extracts spatial features in addition to temporal features from fuzzified input data through a series of modified convolutional layers followed by activation and pooling layers. Here, multilayer CNN is used, and it will take the fuzzified input and apply convolution operation to extract the special features and temporal features. The convolution of the input data with the corresponding filters is expressed as,

$$Z_{i,j}^{(k)}=\int_{k=1}^{10}\left[\sum_{m=1}^{M}W_m^{(k)}\cdot\chi_{i+m-1}+\sum_{n=1}^{N}W_n^{(k)}\cdot\chi_{j+n-1}\right]+b^{(k)}(x) \quad (2)$$

Where $Z_{i,j}^{(k)}$ is represented as the obtained output of the feature map of the k^{th} filter, and then the filter weight for the k^{th} filter at various positions (m,n) is found to be ($W_m^{(k)}$, $W_n^{(k)}$). The input for the convolution layer is represented to be $\chi_{i,j}$.

In addition, with the normal terms, the bias terms are to be considered for the k^{th} filter and it is mentioned as $b^{(k)}(x)$. The output obtained from the convolutional layer is then fed to the activation unit. Here, the ReLU (Rectified Linear Unit) is considered to be the activation function, and it is given as

$$f(x)=\max(0,\tau) \quad (3)$$

Where the pooling layer present in the CNN will reduce the spatial dimensions. Here, Max pooling is used and will focus on the most relevant features for determining the spatial and temporal features.

3.3. BiLSTM for Sequential Analysis

The output from the CNN layers is the extracted features, which will be in both spatial and temporal form and fed into the BiLSTM network. BiLSTM will process the sequential data in both directions (Forward and Backward). The usage of BiLSTM will determine the dependencies in

both directions. Hence, it is more suitable for interpreting the time-series data. Since BiLSTM will consider both the past (forward) and future (backward) information and hence can be suitable for diagnosis and prognosis. Normally, past issues in health may be the cause for the current critical issues, so BiLSTM provides a useful way to determine the long-term dependencies. By using this method, the progression of heart disease based on historical data and short-term fluctuations is analyzed with better understanding. A total of four levels of operations are formulated in the BiLSTM analysis, and they are Input gate (Ig), forget gate (fg), Cell state (Cs), Cell state Update (Cs(u)), Output gate (Og) and Hidden State (Hs). For capturing the long-term dependencies in the overall sequences, the BiLSTM will decide which parts of the previous cell state should be retained or discarded. This was represented by the forget gate (fg) and was mathematically expressed as,

$$f_g = \sigma \sum_{t=1}^{n} (W_f . [h_{t-1}, x_t] + b_f(x)) \quad (4)$$

Where Wf is the weight matrices, and ht is the hidden state approximately taken at time t, and bf is considered to be the bias terms. The Input Gate (Ig) provided in the algorithm will decide which parts of the new information to store in the cell state. This was mathematically expressed as,

$$I_g = \sigma \sum_{t=0}^{n} \sum_{\tau=1}^{m} (W_i . [h_{t-1}, x_t]^{\tau} + b_i(x)) \quad (5)$$

The corresponding sigmoid function is represented as σ. The cell state (Cs) is continuously updated based on the number of iterations done in between the forget gate and the input gate and is expressed as,

$$C_s = \sum_{t=1}^{l} (f_t . C_{s-1} + i_t . \tilde{C}_s) \quad (6)$$

Due to the continuous iteration, the cell state is being updated. To generate this update, the values that were obtained from the current input, as well as the output of the hidden state that came before it, are utilized. This is mathematically expressed as,

$$\tilde{C}_s(u) = \tanh(W_c . [h_{s-1}, x_s] + b_c) \quad (7)$$

Then, the output is determined after continuous iteration, and the main part, which corresponds to the cell state, is processed as the output and is mathematically expressed as,

$$O_g = \sigma \sum_{t=1}^{n} (W_o . [h_{t-1}, x_t] + b_o(x)) \quad (8)$$

The LSTM cell's final output, computed by filtering the cell state through the output gate, is processed after the inclusion of the hidden parameters, and the hidden state is expressed as,

$$H_s = O_g * \tanh(C_s(u)) \quad (9)$$

The BiLSTM works by processing the future content and the past context. Hence, two layers are used for processing both in the forward and backward direction to get the final predicted results. Normally, the patient data can often include sequential information, such as periodic check-ups, test results, or medication changes. The BiLSTM processes this sequence both forward and backwards, enabling it to capture complex temporal relationships that might indicate the progression of heart disease. The forward layer in BiLSTM processes the input sequence $X = [a_1, a_2, ..., a_T]$ from t=1 to t=T, where T is obtained as the total number of time steps, which is mathematically expressed as,

$$\vec{H}_s = \text{BiLSTM}(x_t, \vec{h}_{t-1}) + \text{BiLSTM}(x_t, \vec{h}_{t+1}) \quad (10)$$

The combined output from the forward and backward LSTMs gives a richer understanding of the temporal dependencies, improving prediction accuracy. After the BiLSTM layer, the combined hidden states from both directions are typically passed through a fully connected (dense) layer. The output layer output is expressed as,

$$y(n) = \sigma\left(W_d . H_S^{\text{BiLSTM}} + b_d(x)\right) \quad (11)$$

Where W_d is the weight matrix for the dense layer. H_S^{BiLSTM} is given as the final step output (after continuous iterations), and b_d is the bias term. The BiLSTM can model the sequential nature of medical data, capturing both short-term fluctuations and long-term trends that influence heart disease outcomes. By analyzing both past and future contexts, the model gains a comprehensive understanding of how health indicators evolve over time. Sequential modeling with BiLSTM can uncover subtle patterns in the progression of heart disease that might be missed by traditional models, leading to more accurate predictions.

3.4. Adaptive Stochastic Gradient Descent (ASGD)

The obtained results after BiLSTM are processed again through the fine-tuning process. The fine-tuning of the model is done by using ASGD. It normally adjusts the value of the learning rate dynamically depending upon the gradient history by improving the convergence speed and also prevents overfitting. The weight update rule is mathematically represented as,

$$\sum_{t=1}^{n} \theta_{t+1} = \sum_{t=0}^{1} [\theta_t(x) - \eta_t \theta_t \psi_t R_t(\theta_t) + \beta(x)\chi(t)] \quad (12)$$

Where Θt is the model parameters considered in real-time (i.e., for time t) and ηt is the model learning rate at different time intervals, t. In addition, with this, the loss function is represented as J(θt), and the corresponding gradient in association with the loss function is represented as $\psi_t R(\theta_t)$ with respect to the other parameters, and β(x) is the variation with respect to time. In ASGD, to modify the learning rate, the variance of the gradients is utilized. It is mathematically expressed as,

$$\eta_{t+1} = \sum_{t=0}^{n} \left[\frac{\eta_0}{1+\alpha.t}\right] \quad (13)$$

Where η0 is marked to be the initial learning rate, the decay factor is represented as α, and t is the steps determined during each iteration. The proposed model uses binary cross entropy function as the loss function, and hence it supports optimization function as represented as,

$$L_{ce} = -\frac{1}{N}\sum_{i=1}^{n}[y_i \log_2(Xp_i) + N(x)] + \sum_{j=1}^{m}[(1 - Xy_j) \log_2(1 - Xp_j) + N(x)] \quad (14)$$

Where N is the number of Samples for each interval and is mentioned as N(x), Yi is the true label for the ith sample, when included, X will add real-time feasibilities for prediction, and Pi is the predicted probability for the ith sample. For each mini-batch of training data, perform a forward pass, compute loss, and calculate the gradients through backpropagation. Repeat the process with randomly selected mini-batches to compute gradients and update weights. Reducing the learning rate over time prevents overshooting and helps the model settle in a good solution space.

4. Results and Discussion

Python is used to implement the proposed heart disease prediction model using a combination of an FCNN and BiLSTM. The experimentation is done by using Python 3.6, TensorFlow 2.2.0, Keras 2.3.1, NumPy 1.18.5, Pandas 1.0.5, SciPy 1.4.1, scikit-learn 0.22.2, Matplotlib 3.2.2, Fuzzy Logic Library (scikit-fuzzy 0.4), and CUDA Toolkit 10.1. The development setup consists of Intel Core i7 (10th generation), 64GB with Ubuntu 20.04 operation system. The dataset contains information about the ages and genders of patients, as well as their blood pressure, cholesterol levels, and other clinical markers relevant to cardiovascular disease. There is a binary target variable that indicates whether or not cardiopulmonary disease is present in the individual. The overall dataset is split into training, testing, and validation, as shown in Table 3.The preprocessing stage is crucial for transforming raw data into a clean and structured format ready to be used by the model. The various steps, such as handling missing values, normalization, dimensionality reduction, and fuzzification, are involved in the preprocessing stage. Table 4 illustrates in detail the missing values before and after the model's implication.

Table 3. Dataset splitting

Parameters	Training	Testing	Validation
Age	50	60	80
Height	65	75	103
Weight	68	78	107
Gender	54	64	86
Systolic blood pressure	45	55	73
Diastolic blood pressure	45	55	73
Cholesterol	70	80	110
Glucose	57	67	91
Smoking	20	30	35
Alcohol consumption	30	40	50
Physical activity	10	20	20
Total	514	624	826

Table 4. Missing values (Before and after implication of the proposed system) and normalization results

Features	Missing Values		Normalization	
	Before	After	Original	Normalized
Age	0	0	50	0.65
Height	4	0	65	0.76
Weight	6	0	68	0.76
Gender	2	0	54	0.62
Systolic blood pressure	7	0	45	0.53
Diastolic blood pressure	8	0	45	0.56
Cholesterol	6	0	70	0.74
Glucose	3	0	57	0.65
Smoking	6	0	20	0.34
Alcohol consumption	2	0	30	0.36
Physical activity	4	0	10	0.21

Normalization was performed using Min-Max scaling, transforming all continuous features to a range of [0, 1] to ensure that features are on a similar scale, which aids the model's convergence during training.

Figure 2 shows the distribution of clinical features after normalization, indicating that all features are within the [0, 1] range. Normalization improved the uniformity across features and allowed the model to train more efficiently by ensuring that larger magnitude features did not dominate smaller ones. Then, Fuzzification was applied to introduce uncertainty into the input features, making the model robust to minor variations and imprecision in clinical measurements. The fuzzification process transformed crisp values into fuzzy sets using linguistic variables (e.g., low, medium, high). This is shown in Table 5.

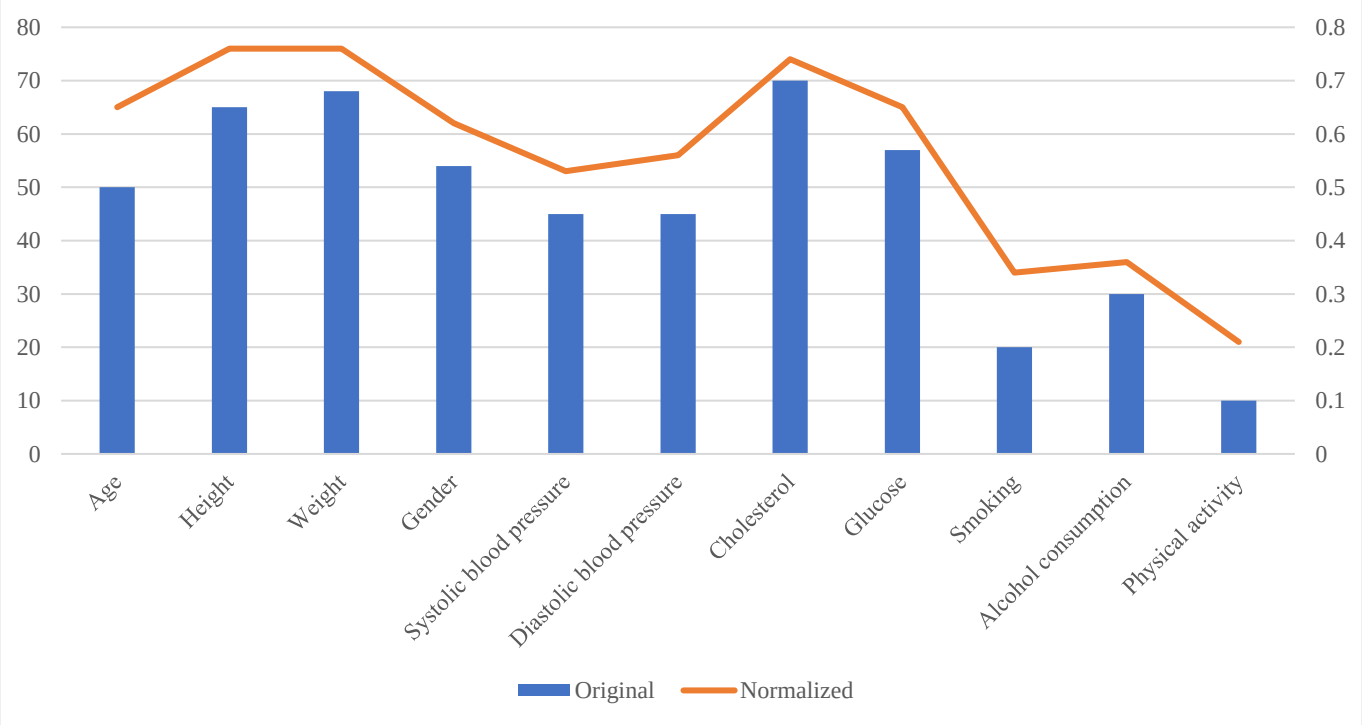

Fig. 2 Graph representing before and after normalization

Table 5. Fuzzified inputs for each feature

Features	Original Value	Fuzzified Values (Low, Medium, High)
Age	50	(0.1,1.0,0.1)
Height	65	(0.2,0.4,1.0)
Weight	68	(0.1,0.7,0.8)
Gender	54	(0.6,0.4,0.6)
Systolic blood pressure	45	(0.2,0.9,0.8)
Diastolic blood pressure	45	(0.4,0.2,0.6)
Cholesterol	70	(0.3,0.5,0.7)
Glucose	57	(0.9,0.7,0.5)
Smoking	20	(0.4,0.5,0.6)
Alcohol consumption	30	(0.6,0.3,0.5)
Physical activity	10	(0.7,0.3,0.8)

Figure 3 shows the fuzzy membership functions for selected features by visualizing how crisp inputs are converted into fuzzy values. Fuzzification added robustness by accounting for uncertainty in clinical measurements, which is common in real-world medical data. This process allowed the MFCNN to learn more generalizable patterns from the input data.

After preprocessing, the class distribution was balanced, which ensures the model does not favor one class over another. The balance was maintained post-preprocessing, with approximately equal representation of heart disease cases (positive class) and non-heart disease cases (negative class). The distribution of the class after preprocessing is illustrated in Figure 4.

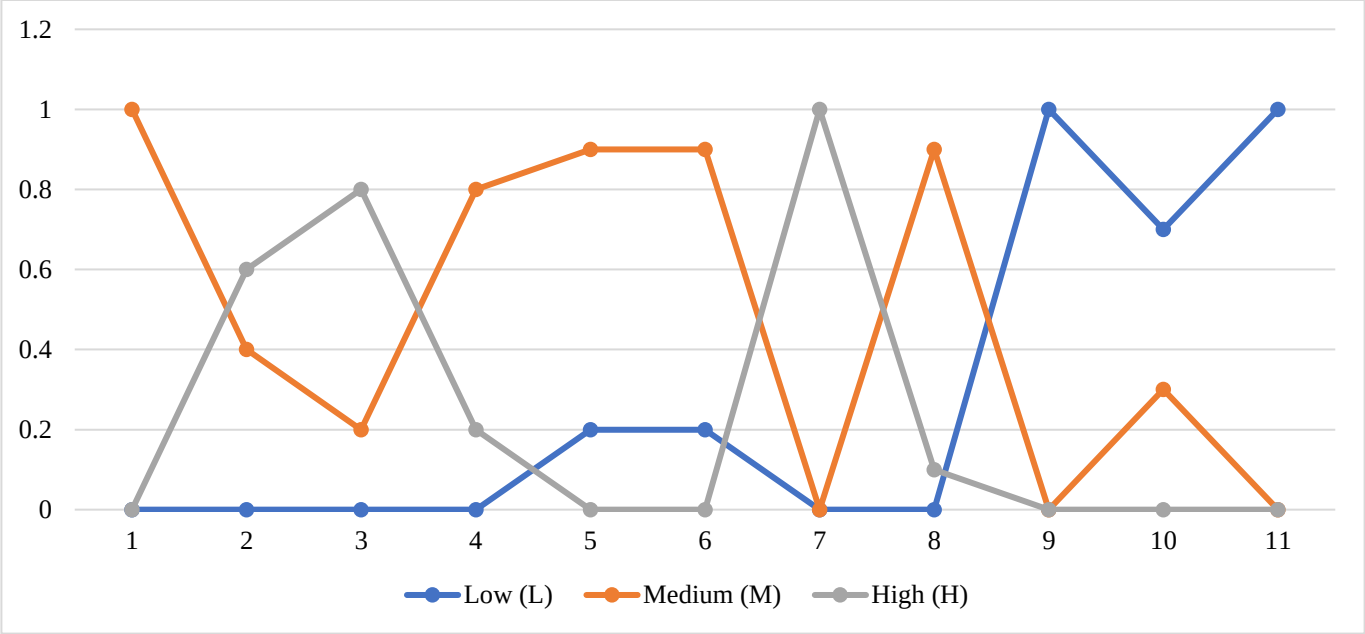

Fig. 3 Fuzzified Input Membership Functions

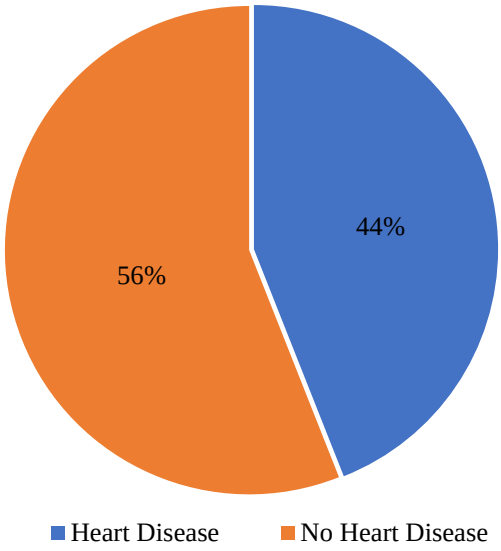

Fig. 4 Class distribution pie chart

Figure 4 shows that the class distribution shows that the dataset remains balanced after preprocessing. The proposed model's performance is analyzed deeply through the path of accuracy over various epochs in both training and validation data. Figure 5 illustrates the accuracy of the proposed model in consideration of the epochs. In Figure 5, the accuracy of the proposed model with respect to validation is displayed over a period of eighty epochs, and the same is also given for training. The epochs are displayed along the x-axis, and the accuracy percentage is displayed along the y-axis of the illustration. Both the training and validation processes are improved in terms of accuracy when the number of epochs is increased. Nevertheless, there is a slight disparity between the two lines, which brings about the possibility that the model is becoming overly accustomed to the data it was trained on. This phenomenon occurs when a model learns the patterns of the training set to the point where it performs poorly on new data. This condition is known as "pattern learning". Evaluation of the training and validation losses of the proposed model is carried out with the help of the features that have been identified, as shown in Figure 6.

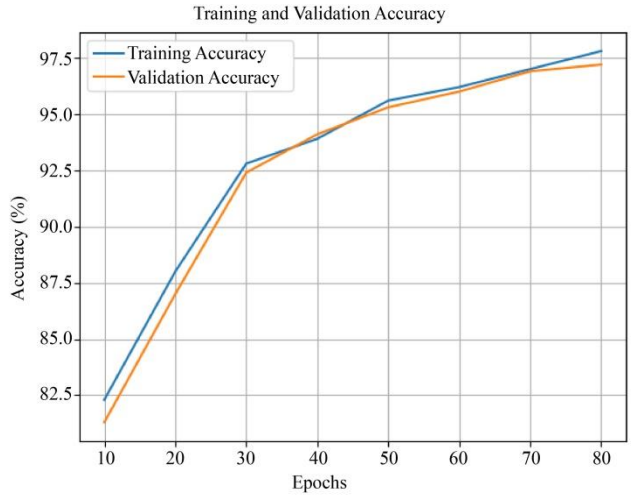

Fig. 5 Training and validation accuracy (models' performance)

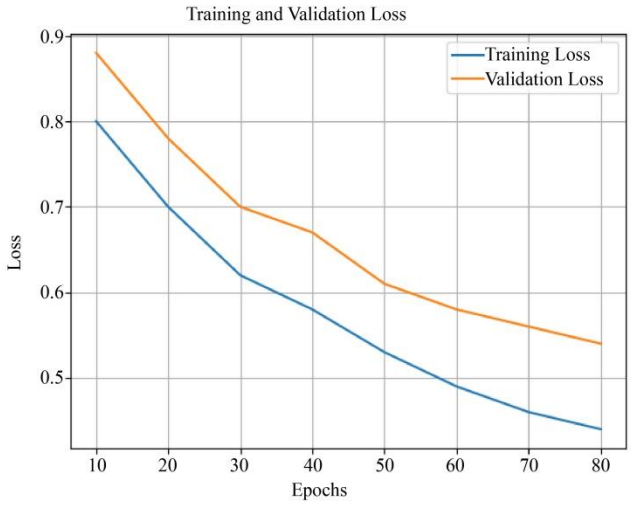

Fig. 6 Training and validation losses

Figure 6 illustrates the loss that the proposed model has experienced over a period of eighty epochs. There is a clear indication that the model is gaining more and more knowledge from the data as each epoch goes by, as evidenced by the decreasing training and validation losses. On the other hand, the training loss continues to decrease, whereas the validation loss begins to level off around the 50th epoch. This leads one to believe that the model may be overfitting, which indicates that it is becoming overly proficient at learning the patterns in the training set and cannot generalize well to new data.

4.1. Confusion Matrix

The performance of the model on the test set is evaluated by the confusion matrix, which takes into account the counts of various sensitive values other than normal values. The values taken are true positive information, true negative value, and false positive value, and confusion occurs over whether it might use these counts. A demonstration of the procedure for generating the confusion matrix for the test set following training is presented in Figure 7. A demonstration of the effectiveness of the proposed model for predicting heart disease is provided by the confusion matrix, which can be found in Figure 7. On one side, we have the actual classes, and on the other side, we have the predicted classes, which are displayed on the x-axis as either positive or negative. Each cell's figure displays the number of instances classified correctly or incorrectly, depending on the situation. According to the cell that contains the intersection of "Actual Positive" and "Predicted Positive," there were 58 instances in which the positive predictions were absolutely accurate predictions. According to the matrix, the model demonstrates an impressive level of accuracy in predicting heart disease, performing a greater number of correct classifications than it does incorrect classifications.

An in-depth analysis of metrics such as recall, precision, and F1-score could be carried out to achieve a more comprehensive understanding of the model's effectiveness, as demonstrated in Table 6. Table 6 provides the evaluation metrics that demonstrate the performance of the Fuzzy-Enhanced CLSTM model in heart disease prediction. The identified precision is marked to be 96.47%, and this brings out most of the possible cases. Then, the accuracy is identified to be 95.17% which is higher when compared to the other traditional methods. Following this, the recall rate is obtained to be 93.97%, and the number of false negatives is reduced compared to the true positives. Finally, the F1-score value is 94.57%, which enhances the evaluation performance of the model. Then, the ROC-AUC score is obtained to be 0.97, which indicates the capability of ranking the instances accurately. Overall, these metrics suggest that the Fuzzy-Enhanced CLSTM model is a promising approach for early detection of heart disease, effectively capturing complex patterns in medical data and handling uncertainties associated with patient health metrics.

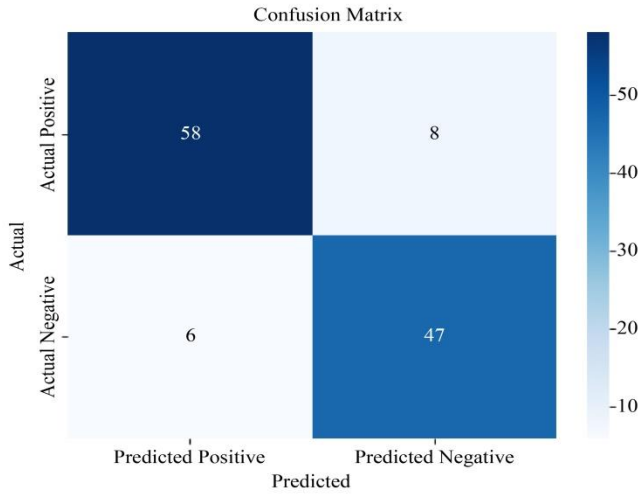

Fig. 7 Confusion matrix

Table 6. Performance metrics of the proposed model

Metric	Prediction Value
Accuracy (%)	96.47
Precision (%)	95.17
Recall (%)	93.97
F1-Score	94.57
ROC-AUC Score	0.97

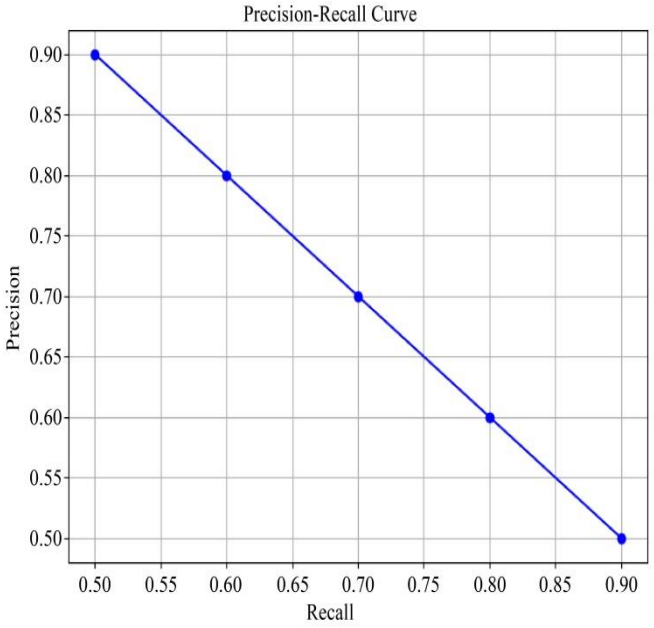

Fig. 8 Precision-recall curve

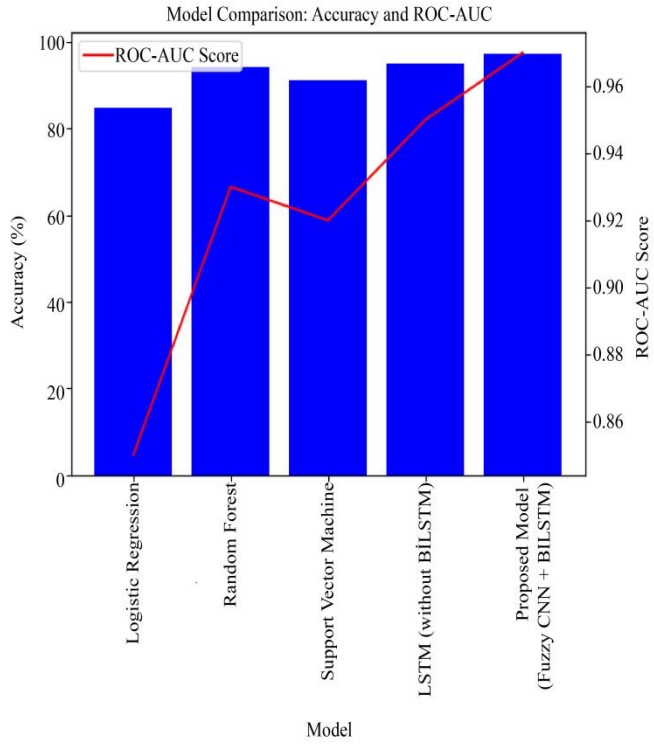

Fig. 9 Comparative performance with baseline models

The precision-recall curve is shown in Figure 8. This curve (Figure 8) helps to visualize the trade-off between precision and recall for various threshold values. A well-performing model typically exhibits high precision and recall simultaneously. Considerably, the proposed model finds a better position in terms of precision and recall. When compared with the traditional machine learning models, the proposed model provides better performance in terms of its performance, as illustrated in Figure 9. The proposed model outperforms other recent methods in detecting and predicting heart disease early and accurately, as shown in Figure 9. The proposed model shows different performances by using fuzzification and without fuzzification. This is illustrated in Figure 10.

The accuracy and ROC-AUC score are evaluated for each model. The results (Figure 10) indicate that the proposed model with fuzzification outperforms the CNN + BiLSTM model without fuzzification, achieving a higher accuracy of 92.5% compared to 89.5% and a higher ROC-AUC score of 0.93 compared to 0.9. This suggests that the incorporation of fuzzification enhances the model's ability to accurately classify instances and discriminate between positive and negative classes. Then, the implication of adaptive stochastic gradient also has better performance, as shown in Figure 11. Figure 11 compares the performance of two optimization algorithms, SGD and ASGD, regarding the number of epochs required to converge and the training time. The obtained experimental results reveal that the ASGD converges faster than SGD, reaching convergence in 76 epochs compared to 110 epochs for SGD. Additionally, ASGD requires significantly less training time, completing its training in 2.1 hours compared to 4.2 hours for SGD.

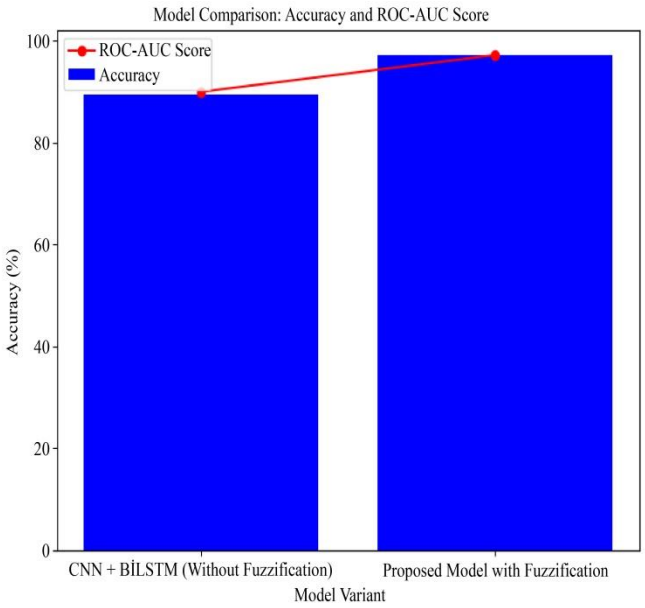

Fig. 10 Effect of fuzzification on accuracy and AUC

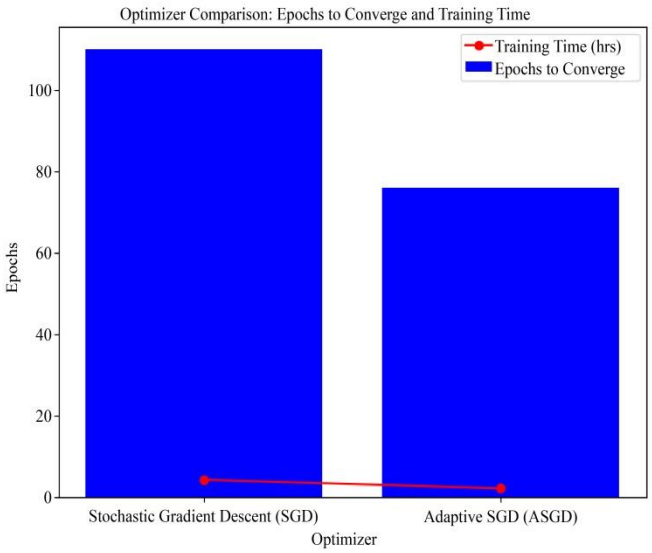

Fig. 11 Training time and epochs for convergence

This suggests that ASGD is a more efficient optimization algorithm for the given model, leading to faster convergence and reduced training time.

5. Conclusion and Future Recommendations

In this study, a novel Fuzzy-Enhanced CLSTM model is proposed that integrates MFCNN along with BiLSTM architecture for heart disease prediction. The main goal of this work was to make predictions more accurate by dealing with the uncertainty and non-linearity that come with patient health data. Including fuzzification in the CNN component made the model more robust in handling imprecise and uncertain data, thereby enhancing its ability to extract critical features from complex clinical parameters. Using BiLSTM's temporal modeling features, the model captured dependencies that changed over time, giving a complete picture of the patient's history. The combination of temporal analysis and spatial feature extraction greatly improves the early and accurate detection of heart disease. By using Adaptive Stochastic Gradient Descent (ASGD), the training speed was increased, and local minima were avoided by letting the learning rate change dynamically. Compared to other traditional methods, this optimization algorithm must make more accurate predictions and work better in more situations. The presented model works better than the other approaches that were tried before, as shown by the test results on datasets that are available to everyone.

The Fuzzy-Enhanced CLSTM model demonstrated enhanced generalization and accuracy, suggesting that it could be a valuable application for the early detection and prediction of heart disease. As a result of its effective training and enhanced sensitivity to critical variations in clinical data, this model exhibits potential for the real-time diagnosis and monitoring of cardiovascular diseases in healthcare settings. Future research will include incorporating suitable classification algorithms to enable the implementation of an efficient healthcare monitoring system for the timely identification of cardiovascular disease. Hyperparameter tuning will moreover be done to maximize the effectiveness of the presented model in predicting a heart attack. The effectiveness of the proposed heart disease prediction framework will additionally be enhanced in the future by more research into meta-heuristic strategies.

References

[1] Xiaoyu Sun et al., "Artificial Intelligence in Cardiovascular Diseases: Diagnostic and Therapeutic Perspectives," *European Journal of Medical Research*, vol. 28, no. 1, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ram Jagannathan et al., "Global Updates on Cardiovascular Disease Mortality Trends and Attribution of Traditional Risk Factors," *Current Diabetes Reports*, vol. 19, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Zeeshan Ahmed et al., "Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine," *Database*, vol. 2020, pp. 1-35, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Fuad Ali Mohammed Al-Yarimi et al., "Feature Optimization by Discrete Weights for Heart Disease Prediction Using Supervised Learning," *Soft Computing*, vol. 25, pp. 1821-1831, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Bharat Singh, Rajesh Kumar, and Vinay Pratap Singh, "Reinforcement Learning in Robotic Applications: A Comprehensive Survey," *Artificial Intelligence Review*, vol. 55, no. 2, pp. 945-990, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Chayakrit Krittanawong et al., "Artificial Intelligence in Precision Cardiovascular Medicine," *Journal of the American College of Cardiology*, vol. 69, no. 21, pp. 2657-2664, 2017. [Google Scholar] [Publisher Link]

[7] Essam H. Houssein, Rehab E. Mohamed, and Abdelmgeid A. Ali, "Heart Disease Risk Factors Detection from Electronic Health Records Using Advanced NLP and Deep Learning Techniques," *Scientific Reports*, vol. 13, no. 1, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Tauhidul Islam et al., "A Systematic Review of Deep Learning Data Augmentation in Medical Imaging: Recent Advances and Future Research Directions," *Healthcare Analytics*, vol. 5, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] William DeGroat et al., "Discovering Biomarkers Associated and Predicting Cardiovascular Disease with High Accuracy Using A Novel Nexus of Machine Learning Techniques for Precision Medicine," *Scientific Reports*, vol. 14, no. 1, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Manju Payal, K. Suresh Kumar, and T. Ananth Kumar, "Recent Advances of Machine Learning Techniques in Biomedicine," *International Journal of Multidisciplinary Research in Science, Engineering and Technology*, vol. 5, no. 5, pp. 772-779, 2022. [Google Scholar]

[11] Paul W. Franks et al., "Precision Medicine for Cardiometabolic Disease: A Framework for Clinical Translation," *The Lancet Diabetes & Endocrinology*, vol. 11, no. 11, pp. 822-835, 2023. [Google Scholar] [Publisher Link]

[12] Fabio Garcea et al., "Data Augmentation for Medical Imaging: A Systematic Literature Review," *Computers in Biology and Medicine*, vol. 152, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Samir S. Yadav, and Shivajirao M. Jadhav, "Detection of Common Risk Factors for Diagnosis of Cardiac Arrhythmia Using Machine Learning Algorithm," *Expert Systems with Applications*, vol. 163, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ehsan Moghadas, Javad Rezazadeh, and Reza Farahbakhsh, "An IoT Patient Monitoring Based on Fog Computing and Data Mining: Cardiac Arrhythmia Use Case," *Internet of Things*, vol. 11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Maggie Yu, Mark Harrison, and Nick Bansback, "Can Prediction Models for Hospital Readmission be Improved by Incorporating Patient-Reported Outcome Measures? A Systematic Review and Narrative Synthesis," *Quality of Life Research*, vol. 33, pp. 1767-1779, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Diego Ortego et al., "Multi-Objective Interpolation Training for Robustness to Label Noise," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6606-6615, 2021. [Google Scholar] [Publisher Link]

[17] Anamika Dhillon, and Gyanendra K. Verma, "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection," *Progress in Artificial Intelligence*, vol. 9, no. 2, pp. 85-112, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] G. Saranya, and A. Pravin, "Hybrid Global Sensitivity Analysis Based Optimal Attribute Selection Using Classification Techniques by Machine Learning Algorithm," *Wireless Personal Communications*, vol. 127, no. 3, pp. 2305-2324, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] T.K. Revathi et al., "An Improved Long Short-Term Memory Algorithm for Cardiovascular Disease Prediction," *Diagnostics*, vol. 14, no. 3, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Aman Darolia et al., "Enhanced Cardiovascular Disease Prediction through Self-Improved Aquila Optimized Feature Selection in Quantum Neural Network & LSTM Model," *Frontiers in Medicine*, vol. 11, 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] David Opeoluwa, Emmanuel Gbenga Dada, and Sanjay Misra, "Diagnosis of Cardiovascular Diseases by Ensemble Optimization Deep Learning Techniques," *International Journal of Healthcare Information Systems and Informatics*, vol. 19, no. 1, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Govindaraj Ramkumar et al., "IoT-Based Patient Monitoring System for Predicting Heart Disease Using Deep Learning," *Measurement*, vol. 218, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] MD Mobin Akhtar et al., "IoMT-Based Smart Healthcare Monitoring System Using Adaptive Wavelet Entropy Deep Feature Fusion and Improved RNN," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 17353-17390, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Manju Payal, T. Ananth Kumar, and K. Suresh Kumar, "Integrating Natural Language Processing (NLP) and Machine Learning Techniques for Healthcare Industries," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 10, no. 5, pp. 3030-3036, 2022. [Google Scholar] [Publisher Link]

[25] Mingle Xu et al., "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," *Pattern Recognition*, vol. 137, pp. 1-12, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Iqbal H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, pp. 1-20, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] Apeksha Shewalkar, Deepika Nyavanandi, and Simone A. Ludwig, "Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235-245, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[28] Qinfeng Zhu, Lei Fan, and Ningxin Weng, "Advancements in Point Cloud Data Augmentation for Deep Learning: A Survey," *Pattern Recognition*, vol. 153, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[29] De Jong Yeong et al., "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review," *Sensors*, vol. 21, no. 6, pp. 1-37, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Vinay Singh et al., "How are Reinforcement Learning and Deep Learning Algorithms Used for Big Data Based Decision Making in Financial Industries-A Review and Research Agenda," *International Journal of Information Management Data Insights*, vol. 2, no. 2, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Fadi AlMahamid, and Katarina Grolinger, "Autonomous Unmanned Aerial Vehicle Navigation Using Reinforcement Learning: A Systematic Review," *Engineering Applications of Artificial Intelligence*, vol. 115, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[32] Luca Bertolaccini et al., "Recent Advances in Lung Cancer Research: Unravelling the Future of Treatment," *Updates in Surgery*, vol. 76, pp. 2129-2104, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[33] Thesath Nanayakkara et al., "Unifying Cardiovascular Modelling with Deep Reinforcement Learning for Uncertainty Aware Control of Sepsis Treatment," *PLOS Digital Health*, vol. 1, no. 2, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] Cristian Drudi et al., "A Reinforcement Learning Model for Optimal Treatment Strategies in Intensive Care: Assessment of the Role of Cardiorespiratory Features," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 5, pp. 806-815, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[35] Smitesh Padte et al., "21st Century Critical Care Medicine: An Overview," *World Journal of Critical Care Medicine*, vol. 13, no. 1, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[36] Jun Ma et al., "Analyzing Driving Factors of Land Values in Urban Scale Based on Big Data and Non-Linear Machine Learning Techniques," *Land Use Policy*, vol. 94, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[37] Jagmohan Kaur, and Baljit S. Khehra, "Fuzzy Logic and Hybrid Based Approaches for the Risk of Heart Disease Detection: State-of-the-Art Review," *Journal of the Institution of Engineers (India): Series B*, vol. 103, pp. 681-697, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[38] L.J. Muhammad, and Ebrahem A. Algehyne, "Fuzzy Based Expert System for Diagnosis of Coronary Artery Disease in Nigeria," *Health and Technology*, vol. 11, pp. 319-329, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] Mario Malcangi, and Giovanni Nano, "Biofeedback: e-Health Prediction Based on Evolving Fuzzy Neural Network and Wearable Technologies," *Evolving Systems*, vol. 12, pp. 645-653, 2021. [CrossRef] [Google Scholar] [Publisher Link]