Baghdad Science Journal

Volume 22 | Issue 9 Article 27

9-23-2025

A Heart Disease Prediction Model Applying Deep Learning and Combined Fuzzy-Mutual Information Based Feature Selection

Annwesha Banerjee Majumder
Department of Information Technology, JIS College of Engineering, Kalyani, India, annwesha.banerjee@gmail.com

Somsubhra Gupta
Department of Computer Science and Engineering, Swami Vivekananda University, Barrackpore, India, gsomsubhra@gmail.com

Sourav Majumder
Department of Cloud Technology, Capgemini India, Kolkata, India, sourav.majumder2016@gmail.com

Dharmpal Singh
Department of Computer Science and Engineering, JIS University, Kolkata, India, dharmpal1982@gmail.com

Sumit Das
Department of Information Technology, JIS College of Engineering, Kalyani, India, sumit.das@jiscollege.ac.in

See next page for additional authors

How to Cite this Article

Majumder, Annwesha Banerjee; Gupta, Somsubhra; Majumder, Sourav; Singh, Dharmpal; Das, Sumit; and Saha, Tanusree (2025) "A Heart Disease Prediction Model Applying Deep Learning and Combined Fuzzy-Mutual Information Based Feature Selection," *Baghdad Science Journal*: Vol. 22: Iss. 9, Article 27.
DOI: https://doi.org/10.21123/2411-7986.5068

A Heart Disease Prediction Model Applying Deep Learning and Combined Fuzzy-Mutual Information Based Feature Selection

Authors

Annwesha Banerjee Majumder, Somsubhra Gupta, Sourav Majumder, Dharmpal Singh, Sumit Das, and Tanusree Saha

BAGHDAD SCIENCE JOURNAL 2025;22(9):3133–3145

RESEARCH ARTICLE

A Heart Disease Prediction Model Applying Deep Learning and Combined Fuzzy-Mutual Information Based Feature Selection

Annwesha Banerjee Majumder [1,*], Somsubhra Gupta [2], Sourav Majumder [3], Dharmpal Singh [4], Sumit Das [1], Tanusree Saha [1]

[1] Department of Information Technology, JIS College of Engineering, Kalyani, India
[2] Department of Computer Science and Engineering, Swami Vivekananda University, Barrackpore, India
[3] Department of Cloud Technology, Capgemini India, Kolkata, India
[4] Department of Computer Science and Engineering, JIS University, Kolkata, India

ABSTRACT

Despite remarkable advancements in medical technology, cardiovascular disease persists as a significant contributor to global mortality. This research addresses the imperative need for timely disease identification through the proposition of a heart disease prediction model, utilizing a merged 1D Convolutional Neural Network (CNN). The main aim of the research is to mitigate the inherent drawbacks of single-layer designs, with a particular emphasis on enhancing hierarchical feature extraction, broadening the model's receptive fields, and facilitating more efficient non-linear transformations. The dataset has been collected from UCI data repository. The study methodology includes a new feature selection strategy that combines Mutual Information and Fuzzy Logic approaches, offering a subtle viewpoint that is not well covered in the literature at the moment. After undergoing extensive training, the output of two 1D CNNs are merged to provide an impressive average accuracy rate of 85%. The integrated 1D CNN model and Explainable AI methodologies yield promising outcomes in heart disease prediction, affirming the model's potential as a screening instrument for early identification and intervention. By providing a strong framework for proactive cardiovascular health care, the study makes a substantial contribution to the nexus between medical science and machine learning.

Keywords: Cardiovascular disease, Convolution neural network, Explainable artificial neural network, Fuzzy logic, Mutual information

Introduction

The development of machine learning has demonstrated its superiority in numerous fields of practical applications. The areas of disease prediction, remote patient monitoring, and drug development are where machine learning excels.[1] Numerous machine learning techniques have already been applied to early cardiovascular disease prediction. According to a World Health Organization research, heart disease is one of the top causes of death, and 9 million people died from it between 2000 and 2019.[2] A significant study topic over the past ten years has been developing novel methods to forecast cardiovascular disease and efficient treatments to slow the disease. The utilization of machine learning techniques has proven to be exceptionally effective in a variety of healthcare areas, which has prompted research into the potential applications of these approaches for the early detection of cardiovascular illnesses. Even though previous studies have made progress in using machine learning to predict CVD, there is still a need for more accurate,

---

Received 26 October 2023; revised 13 October 2024; accepted 15 October 2024.
Available online 23 September 2025

* Corresponding author.
E-mail addresses: annwesha.banerjee@gmail.com (A. B. Majumder), gsomsubhra@gmail.com (S. Gupta), sourav.majumder2016@gmail.com (S. Majumder), dharmpal1982@gmail.com (D. Singh), sumit.das@jiscollege.ac.in (S. Das), tanusree.saha@jiscollege.ac.in (T. Saha).

https://doi.org/10.21123/2411-7986.5068

early detection, and interpretable models. In the light of these circumstances, the study acknowledges the need for a novel strategy that will improve prediction accuracy while also enabling early intervention to lessen the serious effects of heart disease.

The problem addressed in this research is the early and accurate prediction of heart disease, a critical health issue given that heart disease is a leading cause of global mortality. Despite advancements in machine learning, there is a need for more effective models in disease prediction, specifically in the context of cardiovascular diseases. The objective is to create a predictive model that can identify the likelihood of heart disease in individuals at an early stage. The intention is to facilitate early intervention and treatment, thereby preventing the progression of heart disease to more severe and potentially life-threatening conditions. Early prediction is sought to be enabled, allowing for timely medical attention and contributing to better patient outcomes, with the potential to reduce the severity and complications associated with heart disease.

In this paper a heart disease prediction model applying 1D merged convolutional neural network has been proposed. The dataset has been collected from UCI data repository. For feature selection a combined methodology has been applied utilizing Fuzzy Logic and Mutual Information. The integration of fuzzy logic and mutual information in the context of feature selection within data analysis and machine learning presents a flexible and adaptive methodology. Fuzzy logic handles data mistakes and incorporates subject expertise through feature similarity. Mutual information is used in data analysis to evaluate feature significance to the goal variable. Its main goal is to reduce overfitting and streamline high-dimensional datasets. It ensures that the chosen traits are relevant to the situation. Two 1D CNN models have merged together and trained with dataset to achieve enhanced performance. Model interpretability in the healthcare domain is very important. To address this, Shapley Additive Explanations (SHAP) and Local Interpretable Model agnostic Explanations (LIME) frameworks have been employed, which quantifies the contribution of each feature to the model's prediction. This allows us to provide meaningful explanations for the model's decision-making process and identify the key features influencing the predictions. Convolution Neural network is a part of deep learning which is mainly used on image, but can be applied over text data also. The key idea behind CNNs is to leverage the spatial structure and local patterns present in visual data. Unlike traditional neural networks, CNNs are designed to automatically learn hierarchical representations from the input data, capturing features at different levels of abstraction. The use of a merged 1D CNN model provides advantages in terms of enhanced feature extraction, improved model complexity, increased prediction accuracy, robustness to varied input patterns, comprehensive learning, and potential for transfer learning, making it a promising approach for cardiac disease prediction. Explainable Artificial Intelligence is a collection of techniques that enable people to understand the decisions made by machines. LIME proposed by Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrinis[3] is agonistic as it can give explanation for any given supervised learning model and it is local as explains locally faithful within the surroundings or vicinity of the observation. Based on the best Shapley values according to game theory, SHAP (Shapley Additive Explanations) presented by Lundberg and Lee[4] explain individual predictions. The purpose of SHAP is to offer a consistent and equitable method of attributing a model's prediction to its input attributes. The average marginal contribution of each feature over all potential feature combinations is computed, taking into consideration all possible feature combinations. The importance values are guaranteed to satisfy this method's requirements for consistency, symmetry.

The paper aims to address the problem of predicting cardiac illness through the introduction of an innovative model. Specifically, it proposes a 1D merged Convolutional Neural Network (CNN) for heart disease prediction. The key problem being tackled is the development of an effective and accurate predictive model for cardiac diseases using a combination of advanced neural network architecture and a unique feature selection methodology that integrates Fuzzy Logic and Mutual Information. A neural network's ability to combine many 1D CNN layers solves a number of issues with single-layer designs. An individual 1D CNN layer may be unable to perform complex non-linear transformations, have a small receptive field, and have difficulty capturing hierarchical patterns. Multiple layer stacking gives the network the capacity to apply non-linear activation functions at different abstraction levels, learn hierarchical representations, and detect long-range dependencies. Apply merged 1D CNN the performance of the model has been increased by 2.4% on an average. Additionally, the paper focuses on enhancing model interpretability in the healthcare domain by employing Explainable Artificial Intelligence (XAI) methodologies such as Shapley Additive Explanations (SHAP) and Local Interpretable Model agnostic Explanations (LIME).

This study contributes to the field of healthcare analytics by addressing the essential need for accurate and early prediction of cardiovascular disorders through the integration of cutting-edge neural network design, creative feature selection techniques, and XAI This all-encompassing strategy not only contributes in healthcare analytics but also has the ability to transform global health outcomes by enhancing heart disease early detection and intervention tactics.

In next section few significant related worked have been analyzed. After that proposed model detailing and result analysis have been presented. Next conclusion statement has been mentioned.

Related work

Santhana Krishnan J. et al. proposed a model to predict cardiovascular disease using Decision Tree and Naïve Bayes and also analyzed the performance of both.[5] In Reference[6] a heart disease prediction model was proposed using supervised learning mechanism – K Nearest Neighbour, Naïve Bayes, Decision Tree and Random Forest and KNN achieved highest accuracy. A strategy to predict heart disease was proposed by Nichenametla et al. utilizing Naïve Bayes and Decision Tree, where Naïve Bayes performed better on short datasets while Decision Tree performed better on large datasets.[7] An early and accurate detection of cardiac disease was proposed by Muhammad Y et al.[8] with combination of various machine learning approaches like K Nearest Neighbour, Decision Tree, Logistic Regression, Random Forest, Support Vector machine etc. In the article by Jindal H et al. another heart disease prediction model was proposed where Logistic Regression and K Nearest Neighbour were used.[9] A machine learning-based cardiovascular disease prediction model was put forth in the article by Sharma V et al., with the highest results coming from Random Forest over Naive Bayes, Decision Trees, and Support.[10] Rajdhan A. et al. proposed a model to predict cardiovascular disease using machine learning methodology. Naive Bayes, Decision Tree, Logistic Regression and Random Forest were used in the model.[11] A hybrid model for heart disease prediction was proposed by Amin UlHaq et al. where Logistic Regression, K Nearest Neighbor, Support Vector Machine, Naïve Bayes and Decision Tree were used.[12] Raut S et al. proposed heart disease prediction model using different machine learning methods. The methods used in the model were Support Vector Machine, Decision Tree, Naive Bayes and Logistic Regression.[13] P.Jayapratha et al. proposed an efficient model to predict heart disease using traditional machine learning mechanism working on a generated surrogated dataset from UCI dataset and achieved better accuracy than other available models.[14] Patel AC et al. proposed a model for heart disease prediction using KNN, Decision Tree, Logistic Regression, Naïve Bayes, SVM and Random Forest.[15] A model using Logistic regression, Decision tree, SVM, and Naive Bayes was proposed by Islam, Saiful et al. to predict heart disease where 86.25%.[16] Mohan Senthilkumar et al. proposed a hybrid model to predict heart disease using a hybrid mechanism. Random Forest, Neural Network, Decision Tree, Support Vector Machine were used in the model.[17] P. Ramprakash et al. proposed a model for heart disease prediction using Deep Neural Network and S2 statistics with elimination of over fitting and under fitting.[18] Syed Nawaz Pasha et al. proposed a model for predicting cardiovascular disease where they identified that accuracy level achieved through Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Decision Trees (DT) were not so good so they worked on Artificial Neural Network and achieved a good accuracy level.[19] Based on Logistic Regression, K-Nearest Neighbors (KNN), Naïve Bayes (NB), Support Vector Machine (SVM), Neural Networks (NN), and CNN, a model was proposed by T.K Srija et al.[20] Another Deep Learning based model was proposed. Deep Convolutional Neural Network was used with 97% accuracy.[21] Viren Viraj Shankar et al. worked on real hospital dataset and used Convolutional Neural Network to predict heart disease. Accuracy of the proposed model was in between 85% and 88%.[22] Kareem AK, AL-Ani MM, and Nafea AA proposed a 1-D CNN-based method for the diagnosis of autism spectrum disorder based on three different datasets. In terms of accuracy, CNN performs better than any other machine learning model. Adults, children, and adolescents had the greatest recorded accuracy rates—99.45%, 98.66%, and 90%, respectively.[23]

Materials and methods

The proposed model has been built with Convolutional Neural Network and trained with the dataset collected from UCI. SHAP and LIME have been applied over the model to understand the model functionality and decision-making mechanism. Model's block diagram has shown in the Fig. 1 below.

The block diagram of the proposed model has shown in Fig. 1 and detailed model diagram has shown in Fig. 2.

Dataset description

Dataset used in the proposed model has been collected from UCI data repository. The heart disease data set is consisting of 14 features. The histogram of the dataset is shown in Fig. 3.

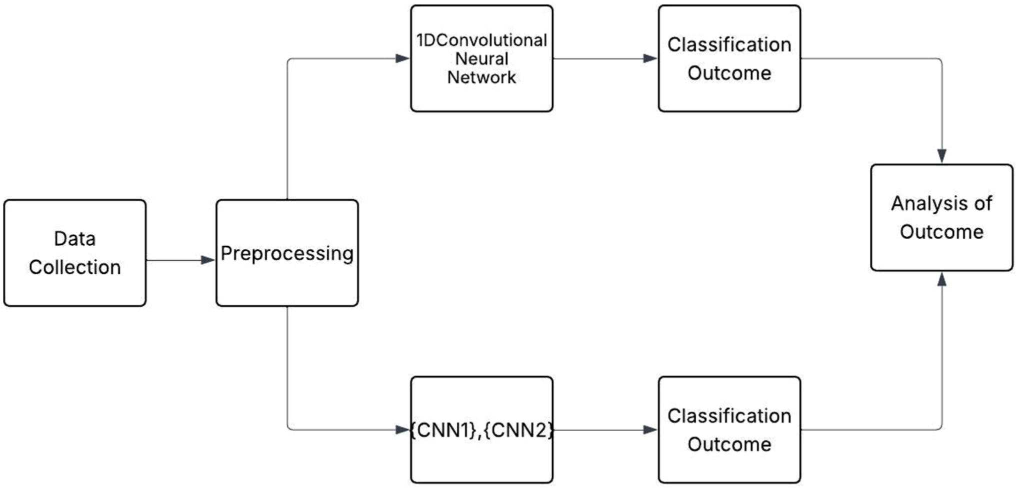


Fig. 1. Proposed model block diagram.

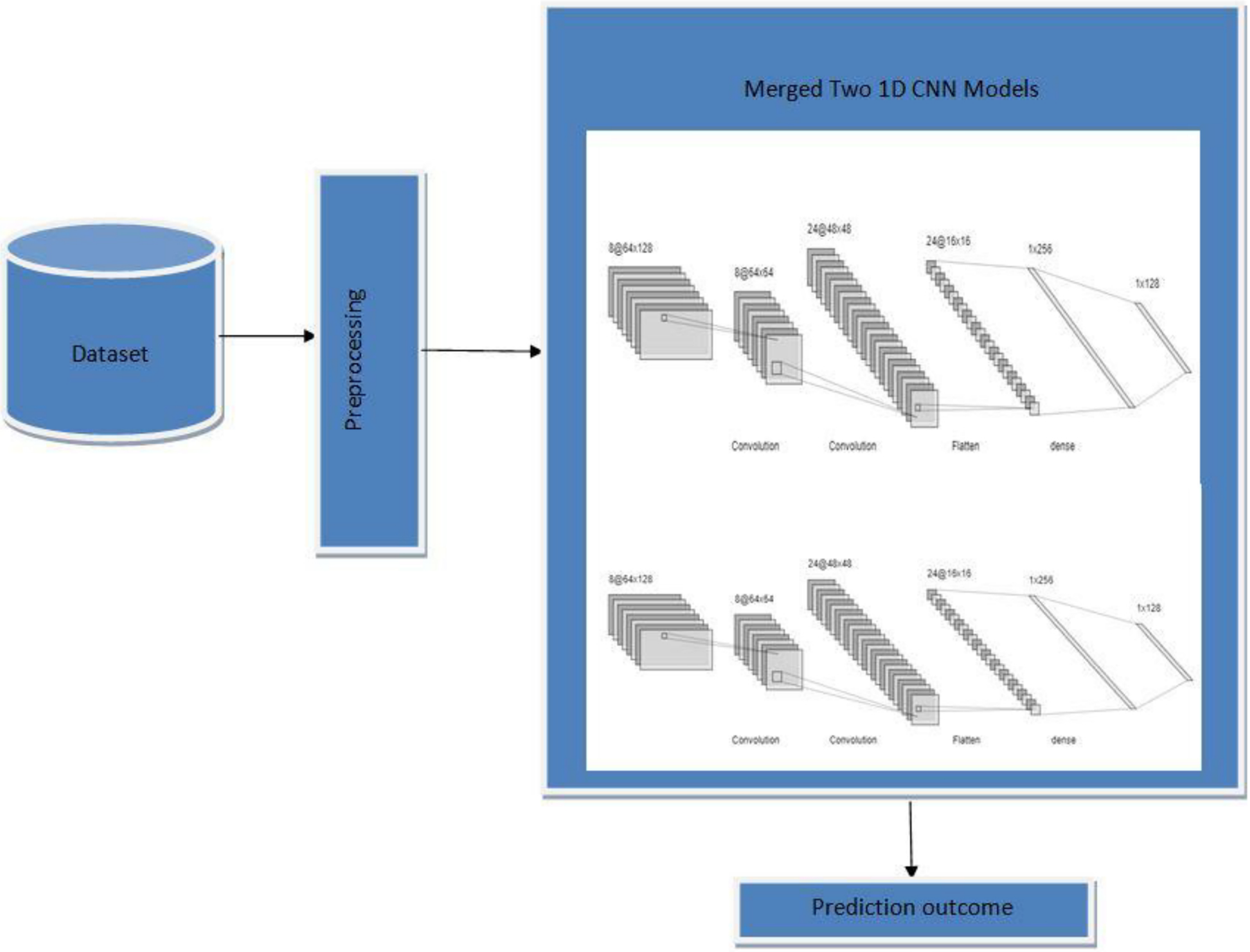


Fig. 2. Detail diagram of proposed model.

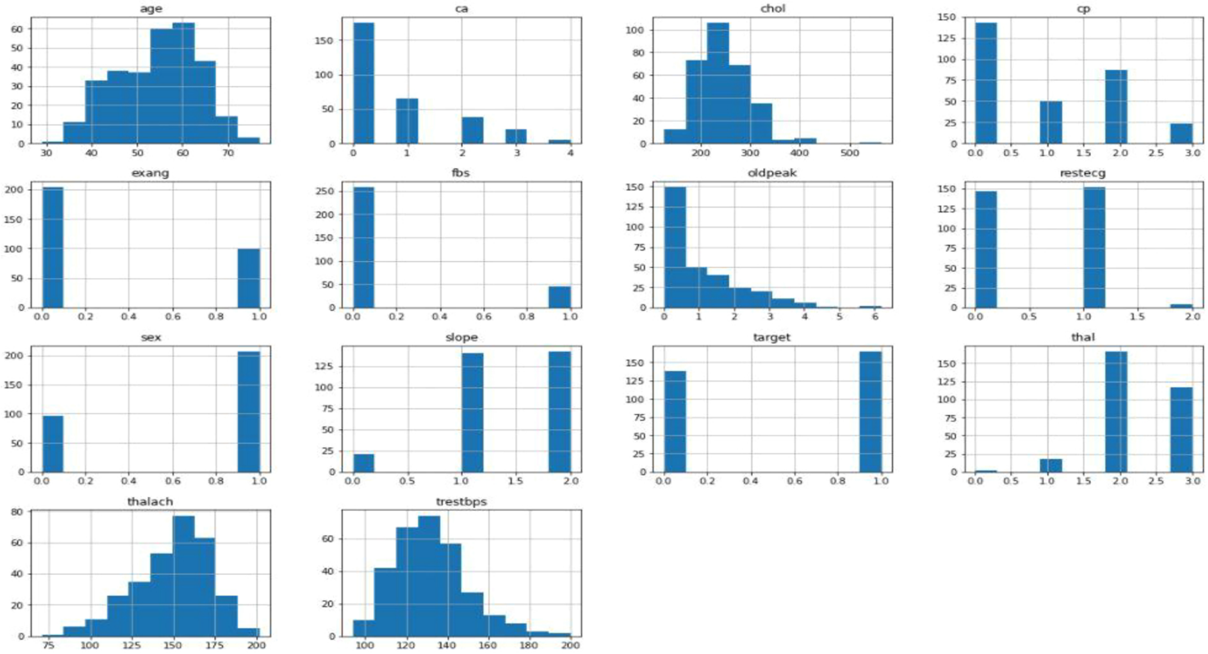


Fig. 3. Histogram of heart disease dataset collected from UCI data repository.

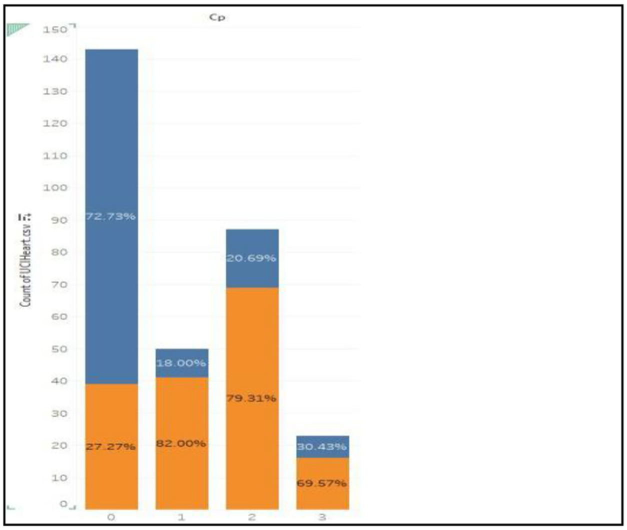


Fig. 4. Chestpain distribution.

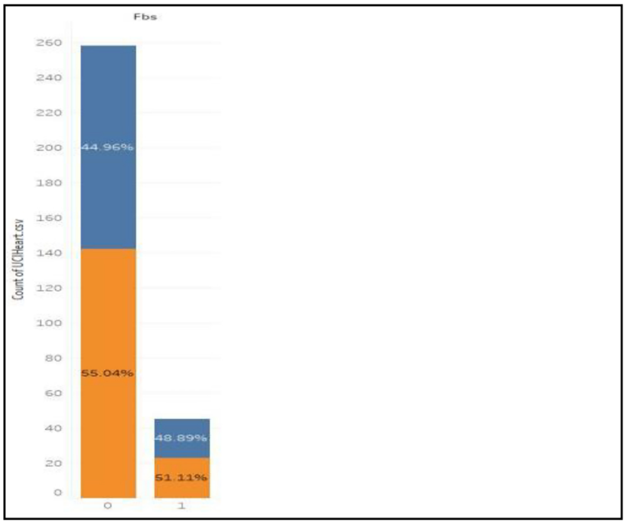


Fig. 5. Blood sugar distribution.

In the following figure, two sample distributions has been shown. Chest pain type distribution over the dataset and Blood sugar distribution over the dataset have been depicted in Figs. 4 and 5.

Feature selection

Feature selection is a critical process in machine learning and data analysis that involves choosing a subset of relevant features from a larger set of available features. It serves several important purposes, including improving model performance, reducing overfitting, enhancing interpretability, and decreasing computational complexity. By selecting the most informative features, we can focus on the essential aspects of the data, leading to more efficient and effective models.

In this proposed methodology, a unique feature selection mechanism combining Fuzzy Logic and Mutual Information has been applied which is yet to circulate in literature widely. Mutual information scores for each feature have been computed, and features have been picked according to a fuzzy logic similarity threshold.

Consider the dataset “Data” and Mutual_ Information(j,y).

Where y is the dependent variable vector and j is certain feature.

Let y represents the binary class labels, where y is a vector of length m.

To ascertain the degree of similarity between a specified feature name and a target feature name, it is essential to calculate a fuzzy similarity score (fscore) for each feature j. The fuzzy similarity function is denoted as Sim_Fuzzy(j).

A fuzzy threshold (T) is a predetermined value that functions as a criterion for feature selection by defining the minimum similarity score necessary. The chosen features are those with Sim_Fuzzy(j) values that meet or exceed the threshold value T.

The feature indices labeled as “fsimilarity” indicate the features with the highest mutual information (MI) scores.

The feature indices that have been chosen are determined by the intersection of fscore and fsimilarity denoted as

$$S_1 = f\,score \cap f\,similarity \tag{1}$$

The set S_1 is defined as the collection of elements i for which the fuzzy similarity Sim_Fuzzy(j) is larger than or equal to the threshold T, with j belonging to the set fsimilarity, as indicated by Eq. (2) below.

$$Sim\_Fuzzy(j) \geq T \tag{2}$$

The features selected applying this mechanism include sex, cp, chol, fbs, thalach, exang, oldpeak, slope, ca, and thal. The histogram has shown in Fig. 6.

Classification

In this phase 1D CNN has been applied for classification of heart disease. The application of 1D CNN enhances the classification accuracy of heart disease cases, leading to improved diagnostic and treatment decisions.

Applying single 1D convolutional neural network model

1D CNN has also been used in the experiment to predict cardiovascular disease. 1D Convolutional Neural Network has been implemented with Keras framework. 1D Convolutional network is a modified version of 2D Neural Network which can be represented with below mentioned Eq. (3).[24]

$$f(x) = \begin{cases} \sum_{i=0}^{k} x(n+i)h(i) & if\ n = 0 \\ \sum_{i-0}^{k} x(n+i+(s-1))h(i), & otherwise \end{cases} \tag{3}$$

Where n is the input length to convolutional layer and x is the input to convolutional layer, k is kernel and h is kernel length.

Activation function used in the proposed model is Rectified Linear Unit (ReLU). Activation functions play a crucial role in generating the output of a neural network by introducing non-linearity. ReLU

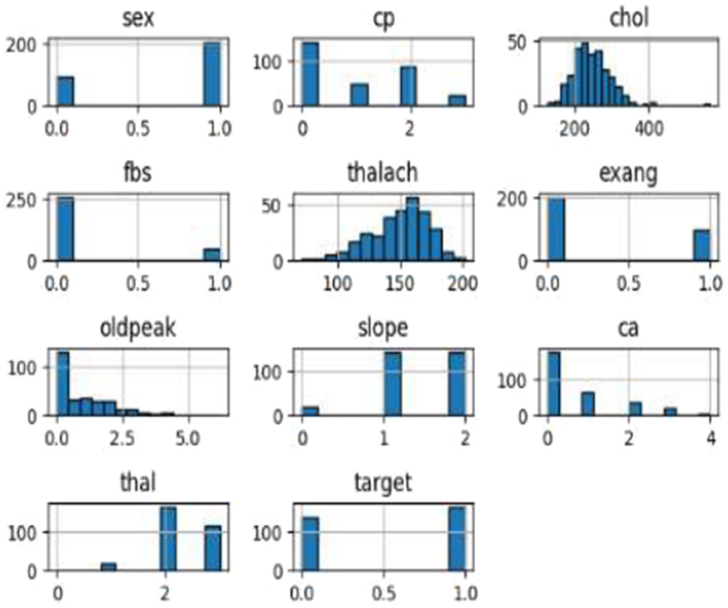
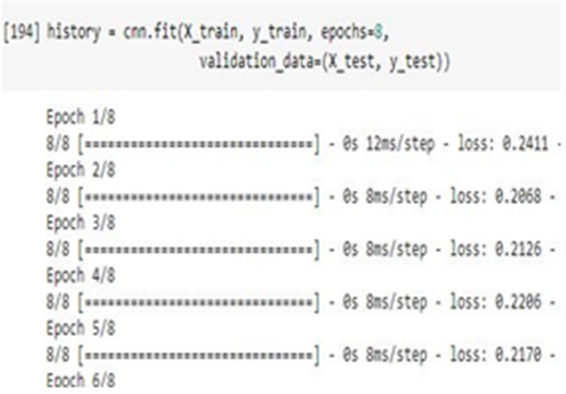


Fig. 6. Histogram of dataset after applying combined fuzzy logic-mutual information for feature selection.



```
[194] history = cnn.fit(X_train, y_train, epochs=8,
                        validation_data=(X_test, y_test))

      Epoch 1/8
      8/8 [==============================] - 0s 12ms/step - loss: 0.2411 -
      Epoch 2/8
      8/8 [==============================] - 0s 8ms/step - loss: 0.2068 -
      Epoch 3/8
      8/8 [==============================] - 0s 8ms/step - loss: 0.2126 -
      Epoch 4/8
      8/8 [==============================] - 0s 8ms/step - loss: 0.2206 -
      Epoch 5/8
      8/8 [==============================] - 0s 8ms/step - loss: 0.2170 -
      Eooch 6/8
```

Fig. 7. Training 1D convolutional neural network.

is considered one of the most efficient activation functions in various types of neural networks. It offers simplicity, computational efficiency, and effective handling of gradient vanishing/exploding problems. By employing ReLU as the activation function, the proposed model aims to enhance the network's representational power and overall performance in capturing complex patterns and relationships within the heart disease data.

It can be represented by following Eq. (4).

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{4}$$

The activation function used in these layers is ReLU. A Flatten layer is used after the convolutional layers to convert the output into a 1D vector. A dense layer is used after the Flatten layer with the activation function ReLU. Performance of the proposed 1D CNN model in each epoch is shown in the Fig. 7.

Another 1D CNN model is built separately with a similar architecture as described above.

Applying merged 1D convolutional neural network model

In this phase two 1D Convolutional neural networks have merged together and trained.

This merging procedure enables the extraction of richer and more significant representations, which may enhance performance in heart disease prediction. To improve its predictive powers and produce precise predictions based on the input data, the combined1D CNN architecture has been trained using the appropriate optimization techniques.

The outputs of the two 1D CNN models have been merged together, possibly using merge operation. This allows the model to combine the learned features from both. 1D CNN models. After merging the outputs, a fully connected dense layer has been applied for the classification of heart disease. The activation function used in this layer is Softmax. Softmax function is commonly used for multi-class classification problems as it provides the probability distribution over the different classes. Adam optimizer has been used during the training of the model with batch size-0.001. Adam is an optimization algorithm that combines the benefits of both AdaGrad and RMSProp, making it suitable for a wide range of deep learning tasks. Binary cross entropy has beeb used as the loss function, which measures the dissimilarity between the predicted probability distribution and the true binary labels.

Model building: two separate 1D CNN models

a: For i = 0 : 2

i: Add two layers of 1D CNN:

$$L_{i1} = f\left(D_{in}; A_{i1}, \ldots, A_{in}\right) || Activation\ function : ReLU \tag{5}$$

Where L_{i1} is the first CNN-1D layer D_{in} is Dataset instance and $A_{i1\ldots n}$ is the feature map for the i^{th} model.

ii: Add a flatten layer.

$$F_i = Flatten\left(X_i : i = 1, \ldots n\right) \tag{6}$$

X_i is the feature set applied to Flatten layer.

iii. Add Dense Layer

$$D_i = Dense\left(x_{inputs}\right) || Activation\ function : ReLU \tag{7}$$

Merged two separate 1D CNN models

$$Merged_{CNN_{fully}conncetd} = Merge(1DCNN : 1DCNN) || Activation\ Function : SoftMax \tag{8}$$

Model Compilation

$$Model = Compile\left(loss : BinarCrossEntropy; optimizer : Adam\right) \tag{9}$$

The summarized description has been shown below.

Layer Conv1D (64 filters, kernel size 3, ReLU activation): This layer applies 64 filters to the input data, each with a kernel size of 3. The ReLU activation function is employed.

Conv1D layer (64 filters, kernel size 3, ReLU activation): The same setup is added to a second Conv1D layer.

Dense layer (100 units, ReLU activation): The ReLU activation function is used in this completely connected layer, which consists of 100 units.

Dense layer (n_outputs units, softmax activation): The softmax activation function is utilized by the final fully connected layer, consisting of n_outputs units.

Results and discussion

This suggested study comprises two distinct experiments. The initial experiment utilized a singular 1D Convolutional Neural Network (CNN) for the prediction of heart disease. This singular 1D CNN architecture has undergone training and evaluation on the dataset to determine its efficacy.

The second experiment was the integration of two 1D Convolutional Neural Networks to form a unified 1D CNN architecture. The integrated architecture has been trained and assessed on the identical dataset to evaluate its performance relative to the standalone 1D CNN model.

The experimental results indicated that the implementation of a singular 1D Convolutional Neural Network (CNN) attained an accuracy range of 80% to 84%. Conversely, the accuracy of the merged 1D Convolutional Neural Network (CNN) architecture varied between 82% and 85%.

Figs. 8 to 11 display the accuracy achieved in each epoch for different executions of the merged 1D CNN. These plots show the variability in accuracy across multiple runs of the merged 1D CNN architecture, highlighting the model's performance stability and consistency.

Fig. 12 represents the performance of Merge 1D CNN models in 100 epochs.

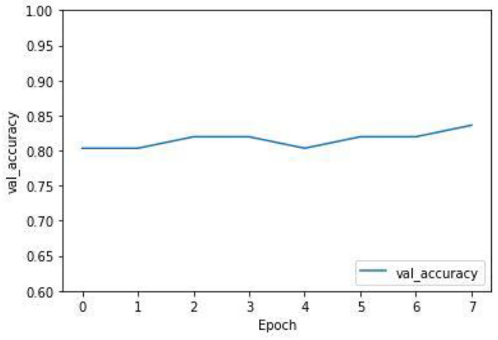


Fig. 8. Performances of single 1D CNN model.

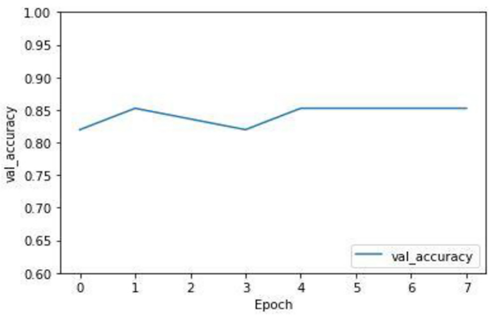


Fig. 9. Performances of merged 1D CNN model.

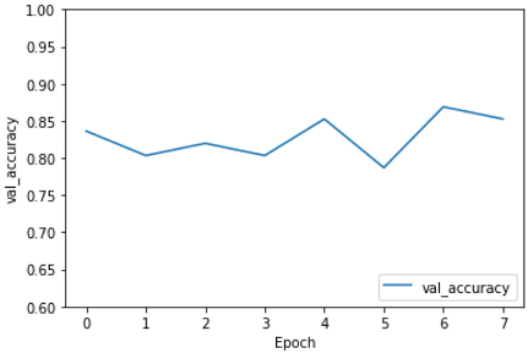


Fig. 10. Performances of merged 1D CNN model.

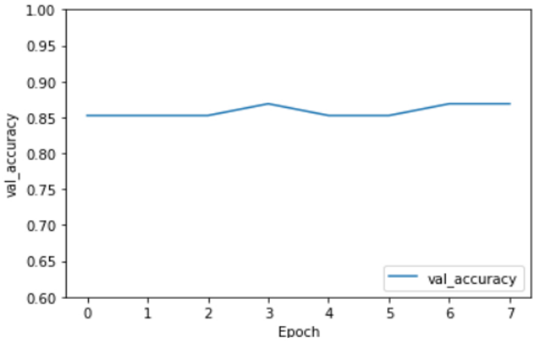


Fig. 11. Performances of merged 1D CNN model.

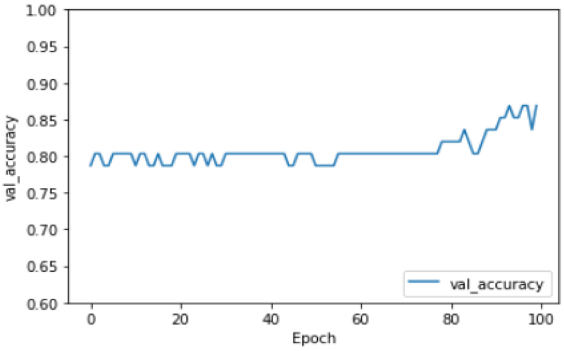


Fig. 12. Performances of merged 1D CNN model (100 epochs).

Fig. 13 below illustrates the resulting confusion matrix from a sample experiment utilizing the merged 1D CNN model.

Along with accuracy other matrices that can be calculated from confusion matrix for model evaluation are precision, recall and F1 score. Precision is a metric that measures the proportion of correctly predicted positive instances (true positives) out of all instances predicted as positive (true positives + false positives).Recall, also known as sensitivity or true positive rate, is a metric that measures the proportion of correctly predicted positive instances (true positives) out of all actual positive instances (true positives + false negatives). The F1 score is a metric that combines both precision and recall into a single value, providing a balanced measure of a classifier's performance. It is the harmonic mean of precision and recall. Performance summary of the proposed merged 1D CNN model has shown in the Table 1 below and which has been visualized in Fig. 14.

In Table 2 the comparative analysis of our proposed models and few existing models has been represented.

The purpose of this analysis is to assess the performance of different models and identify which one performs better in predicting heart disease.

Table 1. 1D merged CNN model performance summary.

Accuracy	85.24%
Precision	85.5%
Recall	82.5%
F1 Score	83.9%

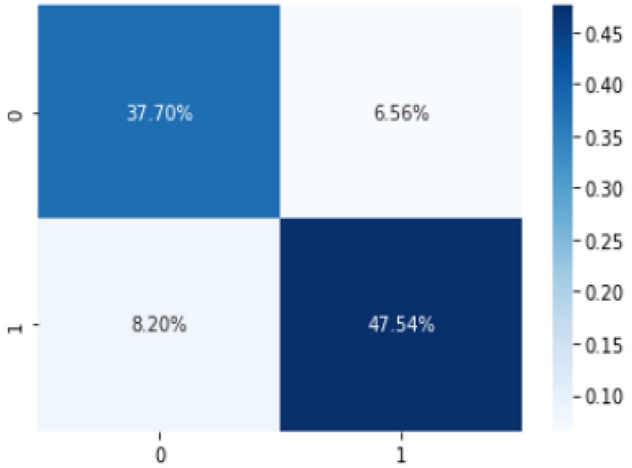


Fig. 13. Confusion matrix of the proposed merged 1D CNN model.

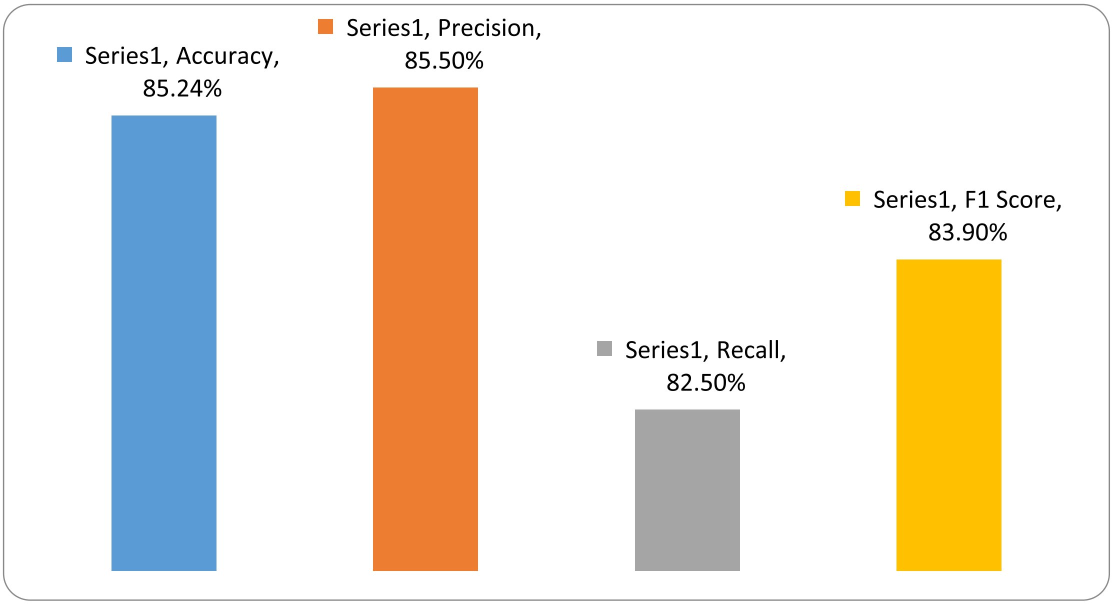


Fig. 14. Model performance summary.

Table 2. Comparative analysis.

Proposed Model	Observations
Ref. [19]	Naïve Bayes and Decision Tree
	Naïve Bayes has given good performance in small dataset
	Decision Tree has given good performance in large dataset
Ref. [20]	Accuracy achieved in different methods :
	Decision Tree : 84%
	Naive Bayes : 83%
Ref. [21]	Logistic Regression : 84%
	KNN(K = 9) : 76%
	Artificial Neural Network : 74%
	SVM (kernel = RBF, C = 100, g = 0.0001) :86%
	SVM (kernel = linear) : 75%
	Naive Bayes : 83%
	Decision Tree : 74%
	Random Forest 83%
Ref. [22]	SVM : 81.75%
	Decision Tree : 80.43%
	Naïve Bayes : 80.43%
	Logistic Regression : 82.89%
Ref. [23]	Decision Tree : 70% (approx.)
	SVM : 52% (approx.)
	KNN : 67% (approx.)
	Random Forest: 78% (approx.)
	Logistic Regression : 84% (approx.)
Our Proposed Model	1D CNN : 80%–84%
	Merged 1D CNN: 82%–85%

The study provides strong evidence that a combined 1D Convolutional Neural Network (CNN) architecture is more effective than a single 1D CNN in predicting cardiac disease.

Table 2 represents comparative analysis of the proposed work with existing works.

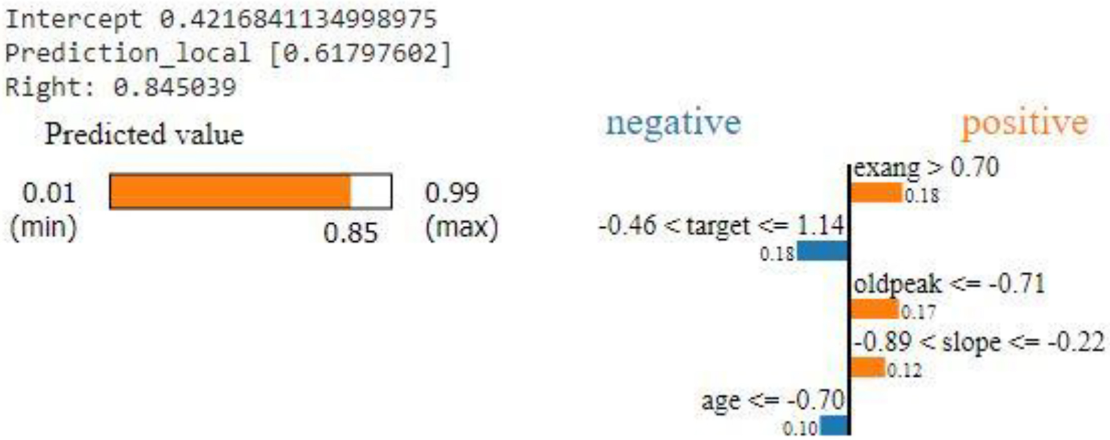
Model explanation with explainable AI

Complex AI models works effectively in different predictive analysis works but the challenges are the difficulty to inherent the models. Mainly in case of disease prediction model expansibility is very crucial for justification of model's performances.

In this proposed work LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations), have applied for model explanations.

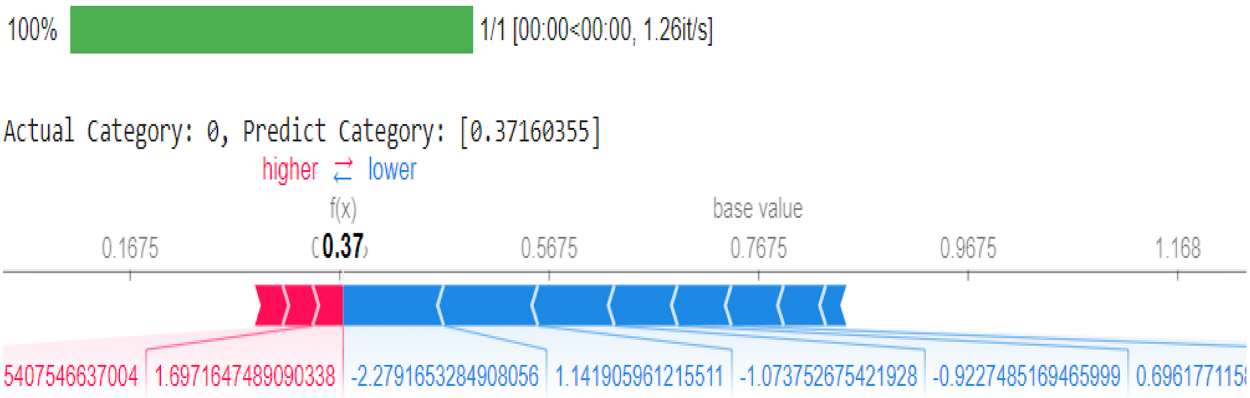
The explanation by LIME has been illustrated in the Fig. 15.

The explanation by SHAP has been illustrated in the Fig. 16.



```
Intercept 0.4216841134998975
Prediction_local [0.61797602]
Right: 0.845039
```

Fig. 15. Model explanation using LIME.



```
100% 1/1 [00:00<00:00, 1.26it/s]

Actual Category: 0, Predict Category: [0.37160355]
```

Fig. 16. Model explanation using SHAP.

Conclusion

In this work a cardiac disease prediction model has been proposed utilizing Merged 1D CNN model. For feature selection combined Fuzzy logic and Mutual Information has been applied. For model explanation LIME and SHAP have been used to have a better understanding of the model. The proposed model has achived an accuracy range 82% to 85%. Additionally, the incorporation of model explanation tactics improves doctors' comprehension of prediction logic, which promotes confidence in the recommendations made by the models. But it's critical to recognize your limitations. Further validation on varied datasets is necessary, as the specific dataset employed may limit the generalizability of our findings. Further research is necessary to address the intricacies of cardiovascular disorders and compare the efficacy of our feature selection process and model explanation methodologies to other approaches, even though they demonstrate promise. However, our work marks a substantial step forward in the application of cutting-edge machine learning methods to improve clinical decision-making and the prediction of cardiovascular disease.

Acknowledgment

We are grateful to our organization to motivate us to carry out the work.

Authors' declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images that are not ours have been included with the necessary permission for republication, which is attached to the manuscript.
- No animal studies are present in the manuscript.
- No human studies are present in the manuscript.
- Ethical Clearance: The project was approved by the local ethical committee at University of JIS College of Engineering.

Authors' contribution statement

AB has analyzed data, design the methodology, computed the experiment, tested the model and written the first draft of the paper. SG has checked the model, edited the manuscript and supervised the work. SM and DS have supervised the overall work.

References

1. Tzanis G, Katakis I, Partalas I, Vlahavas I. Modern applications of machine learning. In Proceedings of the 1st Annual SEERC Doctoral Student Conference-DSC. 2006;1(1):1–10.
2. WHO. Reveals leading causes of death and disability worldwide: 2000–2019.
3. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the predictions of any classifier. arXiv preprint arXiv. 2016;1602.04938. https://doi.org/10.48550/arXiv.1602.04938.
4. Lee SI, Lundberg SM. A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems. December 2017:4768–4777.
5. Santhana Krishan J, Geetha S. Prediction of heart disease using machine learning algorithms. In: 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT). 2019:1–5. https://doi.org/10.1109/ICIICT1.2019.8741465.
6. Shah D, Patel S, Bharti SK. Heart disease prediction using machine learning techniques. SN Comput Sci. 2020;1:345. https://doi.org/10.1007/s42979-020-00365-y.
7. Nichenametla R, Maneesha T, Hafeez S, Krishna H. Prediction of heart disease using machine learning algorithms. Int J Eng Technol (UAE). 2018;7(2.32):363–366. http://dx.doi.org/10.14419/ijet.v7i2.32.15714.

8. Muhammad Y, Tahir M, Hayat M, Chong KT. Early and accurate detection and diagnosis of heart disease using intelligent computational model. Sci Rep. 2020;10:19747. https://doi.org/10.1038/s41598-020-76635-9.
9. Jindal H, Agrawal S, Khera R, Jain R, Nagrath P. Heart disease prediction using machine learning algorithms. IOP Conf Ser.: Mater Sci Eng. 2020:1022. https://doi.org/10.1088/1757-899X/1022/1/012030.
10. Sharma V, Yadav S, Gupta M. Heart disease prediction using machine learning techniques. In: 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). 2020:177–181. https://doi.org/10.1109/ICACCCN51052.2020.9362842.
11. Rajdhan A, Agarwal A, Sai M, Ravi D, Ghuli P. Heart disease prediction using machine learning. Int J Eng Res Technol. 2020;9(04). http://dx.doi.org/10.17577/IJERTV9IS040614.
12. UlHaq A, Li JP, Memon MH, Nazir S, Sun R. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mob Info Sys 2018;2018: 2018(1):1–21. https://doi.org/10.1155/2018/3860146.
13. Raut S, Magar R, Memane R, Rupnar VS. Heart disease prediction using machine learning. Int J Emerging Technol Innov Res. 2020;7(6):2081–2085.
14. Kogilavani SV, Harsitha K, Jayapratha, P, Mirtthula SG. Heart disease prediction using machine learning techniques. Int J Adv Sci Technol. 2020;29(3s):78–87.
15. Patel AC, Mishra M, Shameem A, Chaurasiya S, Saxena A. Prediction of heart disease using machine learning. Int J Sci Dev Res. 2019;4(4). http://dx.doi.org/10.1109/ICECA.2018.8474922.
16. Islam S, Jahan N, Khatun E. Cardiovascular disease forecast using machine learning paradigms. In: 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). 2020. http://doi.org/10.1109/ICCMC48092.2020.ICCMC-00091.
17. Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. IEEE Access. 2019;7:81542–81554. http://doi.org/10.1109/ACCESS.2019.2923707.
18. Ramprakash P, Sarumathi R, Mowriya R, Nithyavishnupriya S. Heart disease prediction using deep neural network. In: 2020 International Conference on Inventive Computation Technologies (ICICT). 2020:666–670. http://doi.org/10.1109/ICICT48043.2020.9112443.
19. Pasha SN, Ramesh D, Mohmmad S, Harshavardhan A, Shabana. Cardiovascular disease prediction using deep learning techniques. IOP Conf Ser.: Mater Sci Eng. 2020;981. http://doi.org/10.1088/1757-899X/981/2/022006.
20. Sajja TK, Kalluri HK. A deep learning method for prediction of cardiovascular disease using convolutional neural network. Rev Intell Artif. 2020;34(5):601–606. http://doi.org/10.18280/ria.340510.
21. Mehmood A., Iqbal M., Mehmood Z. *et al.* Prediction of heart disease using deep convolutional neural networks. Arab J Sci Eng. 2021;46:3409–3422. https://doi.org/10.1007/s13369-020-05105-1.
22. Shankar V, Kumar V, Devagade U, Karanth V, Rohitaksha A. Heart disease prediction using CNN algorithm. SN Comput Sci. 2020;1(170). http://doi.org/10.1007/s42979-020-0097-6.
23. Kareem AK, AL-Ani MM, Nafea AA. Detection of autism spectrum disorder using a 1-Dimensional convolutional neural network. Baghdad Sci J. 2023;20(3(Suppl.):1182. http://doi.org/10.21123/bsj.2023.8564.
24. Srinivasamurthy RS. Understanding 1D convolutional neural networks using multiclass time-varying signals. All Theses. 2018.

نموذج للتنبؤ بأمراض القلب بتطبيق التعلم العميق واختيار الميزات المستندة إلى المعلومات الغامضة والمتبادلة

أنويشا بانيرجي[1]، سومسوبرا غوبتا[2]، سوراف ماجومدير[3]، دارمبال سينغ[4]

[1] قسم تكنولوجيا المعلومات، كلية الهندسة JIS، كالياني، الهند.
[2] قسم علوم وهندسة الحاسوب، جامعة سوامي فيفيكاناندا، باراكبور، الهند.
[3] قسم تكنولوجيا السحابة، كابجيميني الهند، كلكتا، الهند.
[4] قسم علوم وهندسة الحاسوب، كلية الهندسة JIS، كولكاتا، الهند.

المستخلص

على الرغم من التقدم الملحوظ في التكنولوجيا الطبية، إلا أن أمراض القلب والأوعية الدموية لا تزال تساهم بشكل كبير في الوفيات العالمية. يتناول هذا البحث الحاجة الملحة لتحديد المرض في الوقت المناسب من خلال اقتراح نموذج للتنبؤ بأمراض القلب، باستخدام شبكة عصبية تلافيفية مدمجة أحادية الأبعاد . (CNN) الهدف الرئيسي من البحث هو التخفيف من العيوب الكامنة في التصاميم أحادية الطبقة، مع التركيز بشكل خاص على تعزيز استخراج الميزات الهرمية، وتوسيع مجالات تقبل النموذج، وتسهيل التحولات غير الخطية الأكثر كفاءة. تم جمع مجموعة البيانات من مستودع بيانات UCI. تشتمل منهجية الدراسة على استراتيجية جديدة لاختيار الميزات تجمع بين مناهج المعلومات المتبادلة والمنطق الضبابي، مما يوفر وجهة نظر دقيقة لم تتم تغطيتها جيدًا في الأدبيات في الوقت الحالي. بعد الخضوع لتدريب مكثف، تم دمج مخرجات شبكتي CNN أحادية الأبعاد لتوفير معدل دقة متوسط مثير للإعجاب يصل إلى 85%. يوفر نموذج CNN 1D المدمج وأساليب الذكاء الاصطناعي القابلة للشرح نتائج مشجعة في التنبؤ بأمراض القلب، مما يدعم قدرة النموذج على استخدامه كأداة فحص للكشف المبكر والعلاج. ومن خلال توفير إطار قوي للرعاية الصحية الاستباقية للقلب والأوعية الدموية، تقدم الدراسة مساهمة كبيرة في العلاقة بين العلوم الطبية والتعلم الآلي.

الكلمات المفتاحية: أمراض القلب والأوعية الدموية، الشبكة العصبية التلافيفية، الشبكة العصبية الاصطناعية القابلة للتفسير، المنطق الضبابي، المعلومات المتبادلة.