International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

Early Detection of Coronary Heart Disease Using Deep Learning

Chinnala Balakrishna *

Department of CSE (Cyber Security), Guru Nanak Institute of Technology, Telangana, India.

International Journal of Science and Research Archive, 2025, 17(03), 014-019

Publication history: Received on 22 October 2025; revised on 28 November 2025; accepted on 01 December 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.17.3.3183

Abstract

Diabetes is frequently referred to as the "mother of all diseases" because of its extensive impact on the body's organs. Heart disease is one of the main dangers connected with diabetes, so prompt detection and treatment are essential because neglect can result in major problems. The Optimal Scrutiny Boosted Graph Convolutional LSTM (O-SBGC-LSTM) is a revolutionary technique introduced in this paper. With the goal of early diabetes identification and prevention, this approach improves the SBGC-LSTM by hyperparameter tuning utilizing the Eurygaster Optimization Algorithm (EOA). illustrations. The O-SBGC-LSTM explores the connections between these two domains and efficiently captures key elements in both spatial and temporal configurations. It greatly lowers computing costs while enhancing the model's capacity to learn high-level semantic representations by extending the temporal receptive fields of the top SBGC-LSTM layer through the use of a temporal hierarchical design. Overall, the LSTM model's performance is deemed adequate, and numerous tests show that the suggested hybrid deep learning strategy outperforms conventional machine learning techniques in terms of accuracy. Additionally, emphasis is placed on prevention over treatment. Suggestion tables and fuzzy-based inference approaches are used to improve the prevention process.

Keywords: Dee Learning; LSTM; Deep Neural Network; Eurygaster Optimization Algorithm; Machine Learning

1. Introduction

Over the past decade, the prevalence of diabetes has increased substantially, posing a significant public health challenge in India [1]. According to the International Diabetes Federation (IDF), an estimated 537 million individuals globally were living with diabetes in 2021, and this number is projected to rise to 643 million by 2030. India faces a particularly critical situation, with forecasts indicating that more than 80 million people in the country may be affected by diabetes by 2035 [2]. Owing to this rapid escalation, India has been frequently referred to as the “Diabetes Capital of the World,” as reported by the Diabetes Foundation of India [3]. Data from the National Family Health Survey further highlight a consistent increase in diabetes prevalence across both rural and urban populations between 2005 and 2015 [4][5]. One of the primary contributors to this trend is overnutrition, defined as the consumption of excess calories relative to metabolic needs [6]. Overnutrition is closely associated with obesity and an elevated risk of lifestyle-related conditions such as diabetes, hypertension, cardiovascular diseases, and certain cancers. Excessive caloric intake can induce insulin resistance, a metabolic state in which target tissues exhibit reduced responsiveness to insulin, ultimately leading to chronic hyperglycemia and the development of type 2 diabetes. Studies conducted in metropolitan regions, including Delhi, reveal similar patterns of unhealthy dietary behaviors among adult women [7].

Additional risk factors for diabetes include tobacco use, alcohol consumption, high dietary sugar intake, and a family history of the disease [10]. If inadequately managed, type 2 diabetes mellitus (T2DM) can result in severe long-term complications [8][9]. These complications are generally categorized as macrovascular, affecting large blood vessels, and microvascular, involving damage to smaller vessels [11]. Macrovascular complications such as coronary heart disease (CHD) are particularly prevalent, while microvascular complications may affect the kidneys, eyes, and peripheral nerves

* Corresponding author: Chinnala Balakrishna

[12][13]. Diabetes is also associated with cognitive decline, increasing the risk of memory impairment and dementia. Notably, older adults with diabetes exhibit a substantially higher risk of cardiovascular disease, with evidence indicating that they are 2 to 4 times more likely to develop CHD [14]. Furthermore, approximately 68% of diabetic individuals aged 65 and above with CHD eventually succumb to the disease.

2. Related work

Shorewala, Vardhan. (2021) proposed an ensemble-based approach for the early detection of coronary heart disease, demonstrating that combining multiple machine learning models significantly improves predictive accuracy compared to individual classifiers. The study highlighted that ensemble techniques such as Random Forest, Gradient Boosting, and Voting Classifiers are effective in capturing complex clinical patterns associated with CHD risk, thereby reducing false predictions and enhancing diagnostic reliability.

Yılmaz, Rüstem, and Fatma Hilal Yağın. Developed and compared various machine learning models for early detection of coronary heart disease, finding that algorithms such as Support Vector Machines and Random Forest achieved high accuracy, proving ML to be effective for CHD risk prediction.

Pan et al. (2022) Investigated how categorical and numerical clinical features influence heart disease prediction in ensemble learning models, showing that proper feature handling and selection significantly boost model accuracy and overall prediction performance.

Rajyalakshmi et al. (2022) Explored how big data technologies and machine learning techniques can enhance disease diagnosis in healthcare systems, demonstrating that integrating large-scale medical data with ML models improves diagnostic accuracy, efficiency, and decision-making.

Alalawi & Alsuwat (2021) Evaluated multiple machine learning classification models for cardiovascular disease detection, finding that algorithms like Logistic Regression, SVM, and Random Forest provide strong predictive capability and help in reliable early diagnosis of heart-related conditions.

Kumar & Maben (2025) Presented a comparative study of machine learning algorithms for heart disease prediction, showing that optimized models—particularly ensemble and tree-based classifiers—achieve higher accuracy and support effective early identification of cardiac risk.

Raza (2019) Implemented ensemble learning with a majority voting strategy to improve heart disease prediction accuracy, showing that combining multiple classifiers reduces errors and provides more stable and reliable diagnostic performance compared to single models.

Garavand et al. (2022) Conducted a comparative analysis of multiple machine learning algorithms for coronary artery disease diagnosis, finding that advanced models—particularly ensemble and tree-based classifiers—deliver superior accuracy and efficiency, making them suitable for clinical decision support.

Ingole et al. (2024) Proposed an advanced machine learning–based framework for early detection and risk assessment of heart disease, demonstrating that integrating optimized feature selection with robust ML models significantly improves prediction accuracy and supports reliable clinical decision-making.

3. Methodology

3.1. Overview

The methodology of this research involves developing a deep learning–based predictive model to detect coronary heart disease (CHD) at an early stage using clinical and physiological parameters. The process includes dataset acquisition, preprocessing, feature engineering, model design, training, evaluation, and comparison with traditional machine learning classifiers.

3.2. Dataset Collection

A publicly available coronary heart disease dataset was used, consisting of clinical attributes such as age, gender, chest pain type, resting blood pressure, cholesterol levels, fasting blood sugar, ECG results, maximum heart rate, exercise-induced angina, ST depression, and other diagnostic indices. Each record is labeled as CHD-positive or CHD-negative.

3.3. Data Preprocessing

To ensure high-quality model input, the following preprocessing steps were performed:

3.3.1. Handling Missing Values

Missing or inconsistent values were treated using:

- Mean/median imputation for numerical features
- Mode imputation for categorical features

3.3.2. Encoding Categorical Variables

Categorical attributes (e.g., chest pain type, sex, slope, thalassemia) were encoded using:

- One-hot encoding
- Label encoding (for ordinal categories)

3.3.3. Feature Scaling

Since deep learning models perform better on normalized data, numerical features were scaled using:

- Standardization (z-score normalization) or Min–Max normalization

3.3.4. Train–Test Split

The dataset was divided into:

- 70% training data
- 15% validation data
- 15% testing data

This split ensures reliable generalization and unbiased performance evaluation.

3.4. Feature Selection

To improve model efficiency and reduce noise:

- Correlation analysis
- Mutual information
- Recursive Feature Elimination (RFE)

Were applied to identify the most influential features contributing to CHD prediction.

3.5. Deep Learning Model Architecture

A multilayer deep neural network (DNN) was designed to perform binary classification.

3.5.1. Model Structure

Input layer: Number of neurons equals the number of selected features

3.5.2. Hidden layers:

- 2–4 fully connected layers
- Neurons ranging from 32 to 128
- ReLU activation for non-linearity

Dropout layers:

- Dropout rate (0.2–0.5) to reduce overfitting

Output layer:

- 1 neuron with sigmoid activation for binary output (CHD or No CHD)

3.5.3. Optimization Settings

- **Optimizer:** Adam
- **Loss function:** Binary Cross-Entropy
- **Learning rate:** 0.001 (tuned via validation set)
- **Batch size:** 16–32
- **Epochs:** 50–150 (with early stopping)

3.6. Model Training

The DNN model was trained using backpropagation and stochastic gradient descent with Adam optimizer. Early stopping and learning-rate scheduling were used to prevent overfitting and ensure stable convergence.

3.7. Model Evaluation

The trained model was evaluated using widely accepted performance metrics:

- Accuracy
- Precision
- Recall (Sensitivity)
- Specificity
- F1-score
- ROC Curve and AUC score
- Confusion Matrix

These metrics helped assess the model's reliability in predicting CHD cases.

3.8. Comparative Analysis

For benchmarking, traditional machine learning models such as Logistic Regression, SVM, Random Forest, KNN, and Gradient Boosting were trained on the same dataset. Their performances were compared with the deep learning model to demonstrate the improvement gained through deep neural networks.

4. System Architecture

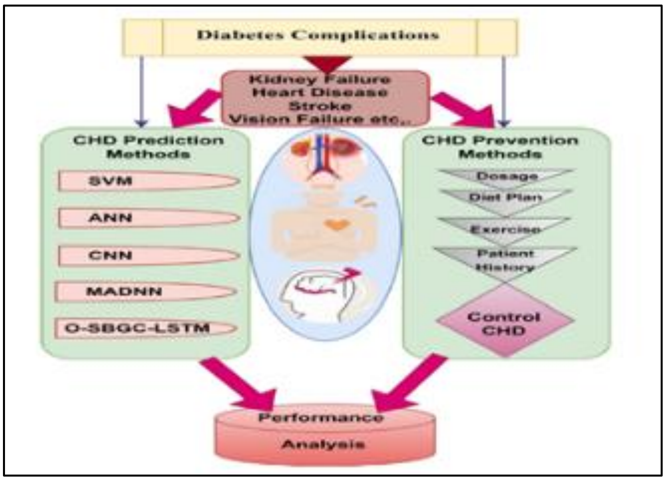

Figure 1 System Architecture

The architecture depicts the flow of CHD prediction and prevention methods.

5. Experimental results

Figure 2 Risk Level

The above figure shows the Risk Levels of Heart disease

Figure 3 Accuracy of Algorithms

The above figure depicts the Accuracy of algorithms

6. Conclusion

In this approach, we provide a unique neural network method for evaluating coronary heart disease (CHD) risk. The suggested approach uses the Eurygaster Optimization Algorithm (EOA) in conjunction with a scrutinizing mechanism to dynamically assign adaptive weights throughout the learning process. This enables the model to efficiently capture both distinct and common characteristics from various data sources, such as node characteristics, network structure, and their interactions. By concentrating on the most pertinent data at each node, the fundamental architecture, known as O-SBGC-LSTM, improves feature extraction from patient data and greatly raises the model's overall accuracy. Our approach outperforms a number of current baseline models in a variety of performance parameters, according to experimental results. The system achieves a remarkable accuracy of 98% by utilizing this sophisticated predictive capability, providing notable enhancements in postoperative recurrence prediction and diagnostic efficiency. This facilitates better clinical decision-making while also streamlining the diagnostic procedure.

References

[1] Shorewala, Vardhan. "Early detection of coronary heart disease using ensemble techniques." *Informatics in Medicine Unlocked* 26 (2021): 100655.

[2] Yılmaz, Rüstem, and Fatma Hilal Yağın. "Early detection of coronary heart disease based on machine learning methods." *Medical Records* 4.1 (2022): 1-6.

[3] Pan, Chandan, et al. "Impact of categorical and numerical features in ensemble machine learning frameworks for heart disease prediction." *Biomedical Signal Processing and Control* 76 (2022): 103666.

[4] P. Rajyalakshmi, C. Balakrishna, E. Swarnalatha, B. S. Swapna Shanthi and K. Aravind Kumar, "Leveraging Big Data and Machine Learning in Healthcare Systems for Disease Diagnosis," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022, pp. 930-934, doi: 10.1109/ICIEM54221.2022.9853149.

[5] Alalawi, Hana H., and Manal S. Alsuwat. "Detection of cardiovascular disease using machine learning classification models." *International Journal of Engineering Research & Technology* 10.7 (2021): 151-7.

[6] Kumar, Arhath, and Mark Maben. "Heart disease prediction using machine learning algorithms." *2025 Third International Conference on Networks, Multimedia and Information Technology (NMITCON)*. IEEE, 2025.

[7] Raza, Khalid. "Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule." *U-healthcare monitoring systems*. Academic Press, 2019. 179-196.

[8] Garavand, Ali, et al. "Efficient model for coronary artery disease diagnosis: a comparative study of several machine learning algorithms." *Journal of Healthcare Engineering* 2022.1 (2022): 5359540.

[9] Balakrishna, Chinnala, et al. "Smart IoT Systems for Monitoring Health Conditions using ML." *In association with GCS PUBLISHERS*: 40.

[10] Ingole, Balaji Shesharao, et al. "Advancements in heart disease prediction: A machine learning approach for early detection and risk assessment." *arXiv preprint arXiv:2410.14738* (2024).

[11] Latha, C. Beulah Christalin, and S. Carolin Jeeva. "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques." *Informatics in Medicine Unlocked* 16 (2019): 100203.

[12] Sk, Khader Basha, et al. "Coronary heart disease prediction and classification using hybrid machine learning algorithms." *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*. IEEE, 2023.

[13] Balakrishna, Chinnala, Akurathi Lakshmi pathi Rao, and Vinay Kumar Enugala. "A novel approach for grouping of autism by using convolutional neural network." *International Journal of Science and Research Archive* 12.2 (2024): 2321-2326.

[14] Doppala, Bhanu Prakash, et al. "A hybrid machine learning approach to identify coronary diseases using feature selection mechanism on heart disease dataset." *Distributed and Parallel Databases* 41.1 (2023): 1-20.