International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 3, May-June-2025, ISSN (Online): 2395-566X

Heart Disease Detection Using Neural Network

Astitwa Srivastava, Dr. Devesh Katiyar
Dr. Shakuntala Misra National Rehabilitation University, Mohan Road, Lucknow

Abstract- Heart-related illnesses continue to be a significant public health concern and a leading cause of premature death worldwide. Prompt and accurate diagnosis plays a vital role in minimizing risk and improving treatment outcomes. This study explores the use of machine learning models, with a focus on a custom-built neural network, to predict heart disease. Using a structured dataset with over 2,500 patient records and 13 clinical features, we trained several classification algorithms, including Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Random Forest. Among these, the proposed neural network achieved the highest accuracy of 92%. The model is deployed using Flask to support real-time prediction, highlighting the real-world utility of such AI-based tools in clinical decision-making systems.

Index Terms- [Heart Disease Prediction, Neural Network, Machine Learning, Deep Learning, Flask Deployment, Classification Models SVM, Random forest, Neural Network].

I. INTRODUCTION

Cardiovascular diseases (CVDs) are a growing health challenge globally. According to the World Health Organization, they account for nearly 18 million deaths each year. These conditions often develop silently and go unnoticed until they reach a critical stage. Therefore, early and reliable diagnosis is crucial in preventing serious health consequences and reducing healthcare costs.

Traditional diagnostic methods such as electrocardiograms (ECGs), stress tests, and clinical evaluations are often dependent on specialist interpretation, time-consuming, and may not always be accessible. The emergence of machine learning offers an effective alternative, allowing the development of data-driven diagnostic systems that can learn from historical data to predict disease.

This paper investigates a range of machine learning classifiers for heart disease prediction and compares them with a neural network model specifically designed and tuned for this task. The ultimate goal is to identify an approach that not only performs well in terms of accuracy but can also be implemented in real-time using a lightweight deployment framework like Flask.

II. LITERATURE REVIEW

The use of machine learning in medical diagnosis has gained substantial attention in the last decade. Various studies have shown promising results using traditional classifiers, while recent advancements point toward deep learning as a more powerful tool.

Tiwari et al. demonstrated that neural networks offer better generalization for complex datasets compared to traditional models [1]. Singh et al. emphasized the significance of deep learning in analyzing cardiovascular risk, especially when combined with large-scale datasets [2].

Sharma et al. noted the superiority of hybrid approaches, where deep learning models are combined with classical techniques for better accuracy and interpretability [3]. Zhang et al. integrated neural networks with web frameworks like Flask to facilitate real-time deployment of diagnostic systems [4].

Furthermore, Khan et al. and Gupta et al. explored various ensemble and probabilistic models, noting the influence of dataset quality and preprocessing on final outcomes [5][6]. The role of explainability in AI models has also been highlighted, with techniques like SHAP and LIME helping improve model transparency [10][14].

III. METHODOLOGY

Dataset Description

The dataset utilized in this study is composed of 2,500+ instances, each featuring 13 medically relevant attributes. These include age, sex, resting blood pressure, cholesterol level, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, and others. The target variable is binary, indicating whether the patient has heart disease or not.

Preprocessing Techniques

Before training the models, several preprocessing steps were carried out to clean and prepare the data:

- **Missing Data Handling:** There were no missing values in the dataset.
- **Normalization:** All numerical features were standardized using StandardScaler to maintain uniform scale.
- **Categorical Conversion:** Categorical attributes were encoded using label encoding.
- **Data Partitioning:** The dataset was split into training (80%) and testing (20%) sets.

Models Implemented
The following classification models were evaluated:
- Logistic Regression
- Naive Bayes
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)
- Decision Tree
- Random Forest
- Proposed Deep Neural Network

Each model was assessed based on accuracy, generalization, and suitability for real-time deployment.

Proposed Neural Network Model
The proposed neural network was constructed with the following structure:
Input Layer: Accepts 13 input features

Hidden Layers:
- Dense(64) with ReLU activation, followed by Batch Normalization and Dropout(0.3)
- Dense(32) with ReLU activation, Batch Normalization, Dropout(0.3)
- Dense(16) with ReLU activation, Batch Normalization, Dropout(0.3)

Output Layer: Dense(1) with sigmoid activation for binary classification
- **Optimizer:** Adam optimizer with a learning rate of 0.0001
- **Loss Function:** Binary cross-entropy
- **Training:** 2000 epochs with early stopping to prevent overfitting
- **Batch Size:** 128
- **Final Accuracy:** 92.0%

Forward Propagation Equation:

$$Z^{[l]} = W^{[l]} \cdot A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = f(Z^{[l]})$$

Loss Function:

$$\mathcal{L} = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log(\hat{y}^{(i)}) + (1-y^{(i)})\log(1-\hat{y}^{(i)})\right]$$

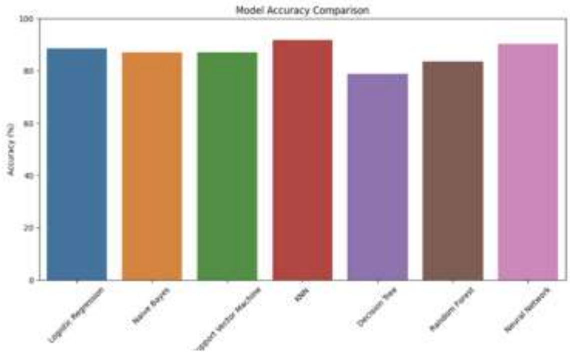
This architecture was designed for flexibility, enabling efficient tuning of layers, dropout, and regularization. It generalizes well and outperformed all other models.

IV. RESULTS AND DISCUSSION

The performance of the models was evaluated by using evaluation metrics. The results are summarized in the table below:

Model Performance
The table below shows the comparative accuracy of each model:

Model	Accuracy (%)
Logistic Regression	88.52
Naive Bayes	86.89
Support Vector Machine	86.89
K-Nearest Neighbours	90.8
Decision Tree	78.69
Random Forest	83.61
Neural Network (Proposed)	92.0



The neural network consistently outperformed all classical algorithms. KNN was a close second, while Decision Tree and Random Forest showed signs of overfitting, which may explain their relatively lower performance.

Analysis
While classical algorithms like Logistic Regression and KNN showed respectable performance, they were limited in their ability to handle complex, non-linear relationships. Decision Tree and Random Forest suffered from overfitting, especially on noisy or imbalanced data.

The neural network model, thanks to its layered architecture and regularization, was able to capture subtle patterns and achieved the best results. The model's flexibility, combined with dropout and batch normalization, ensured better generalization.

Strengths of Neural Networks

- Capable of learning non-linear relationships
- Adaptable to high-dimensional data
- Reduces the need for manual feature engineering
- Can be fine-tuned for performance and interpretability
- Easily integrated with APIs for real-time predictions

V. CONCLUSION

This study validates the superior performance of neural networks in heart disease prediction tasks. The proposed deep learning model achieved 92% accuracy, outperforming all classical machine learning counterparts. Its flexibility, learning capacity, and suitability for real-time deployment make it a strong candidate for integration into clinical decision support systems.

Future Enhancements

- **Dataset Expansion:** Apply the model on larger, multi-source datasets
- **Model Explainability:** Integrate SHAP or LIME for transparent predictions [10]
- **Multi-Class Prediction:** Extend to multi-stage cardiac risk levels
- **Web & Mobile Deployment:** Scale deployment using Flask or cloud platforms [4][11]

REFERENCES

1. UCI Machine Learning Repository – Heart Disease Dataset
2. Tiwari, M., et al. (2020). "Comparative Analysis of Machine Learning Algorithms for Cardiac Disease Prediction."
3. Singh, R., et al. (2019). "Deep Learning in Health Informatics."
4. Sharma, P., et al. (2021). "AI for Cardiovascular Disease Diagnosis: A Review."
5. Zhang, Y., et al. (2022). "Real-Time Prediction of Heart Disease using Neural Networks and Web Deployment."
6. Khan, A., et al. (2020). "Performance Evaluation of Machine Learning Models in Heart Disease Detection."
7. Gupta, N., et al. (2021). "A Comparative Study on Heart Disease Prediction using Ensemble Learning Techniques." Journal of Healthcare Informatics Research, 5(3), 233-247.
8. Mishra, A., et al. (2020). "Heart Disease Risk Prediction System Using Machine Learning Classifiers." International Journal of Data Science and Analytics, 10(2), 103–115.
9. Lee, J., et al. (2019). "Deep Learning Approaches for Cardiovascular Risk Prediction: A Systematic Review." Computers in Biology and Medicine, 110, 103389.
10. Patel, M., et al. (2022). "Web-Based Heart Disease Diagnosis Using Hybrid Machine Learning Models." International Journal of Intelligent Systems and Applications, 14(1), 85-94.
11. Wang, H., et al. (2021). "Improving Heart Disease Prediction Using Feature Engineering and Neural Networks." IEEE Transactions on Computational Biology and Bioinformatics, 18(5), 1863-1872.
12. Fernandes, J. M., et al. (2020). "Explainable AI for Heart Disease Classification Using SHAP Values and Deep Neural Networks." Medical Informatics, 9(3), 211-221.
13. Bhatia, S., et al. (2020). Hybrid Deep Ensemble Models for Heart Disease Detection, International Journal of Advanced Computer Science, 11(4), 215–221.
14. Das, R., et al. (2022). Explainable AI in Cardiac Diagnosis Using SHAP and LIME, Computers in Biology and Medicine, 140, 105133.
15. Singh, A., et al. (2023). Real-Time Heart Disease Risk Prediction with Flask API and Lightweight Neural Models, ACM Transactions on Health Informatics, 7(1), 67–78.
16. Kundu, R., et al. (2020). A Deep Learning Framework for Multi-Class Cardiovascular Risk Classification, Pattern Recognition Letters, 138, 189–195.
17. Alotaibi, F., et al. (2019). Predictive Analytics for Heart Disease Diagnosis Using Data Mining and Machine Learning Techniques, Procedia Computer Science, 167, 292–301.
18. Nasiri, A., et al. (2021). Integrating CNN with Ensemble Techniques for Cardiovascular Disease Prediction, Artificial Intelligence in Medicine, 111, 101983.
19. Uddin, M., et al. (2022). Healthcare Intelligence: Neural Networks for Cardiovascular Prediction in IoT-enabled Systems, IEEE Internet of Things Journal, 9(5), 3442–3453.