International Conference on Intelligent Computing and Explainable AI (ICICEA) – April 2026

RESEARCH ARTICLE OPEN ACCESS

Transparent AI Model For Early Detection Heart Disease

Dr.R.Sudha*, S.Shakina Bevi**, M.Sri Dharshini**, P.Yamini**

* Associate Professor -Department of Computer Science and Engineering , A.V.C. College of Engineering, Mayiladuthurai, India

** UG Students Department of Computer Science and Engineering , A.V.C. College of Engineering, Mayiladuthurai, India

ABSTRACT

Heart disease remains one of the leading causes of mortality worldwide, making early detection crucial for improving patient outcomes. This paper proposes a transparent Artificial Intelligence (AI) model designed to predict the likelihood of heart disease at an early stage. Unlike traditional black-box models, the proposed approach emphasizes interpretability and explainability, enabling healthcare professionals to understand the decision-making process. The model utilizes machine learning algorithms trained on clinical datasets containing patient health parameters such as age, blood pressure, cholesterol levels, and heart rate. Feature importance techniques and explainable AI (XAI) methods are incorporated to provide clear insights into predictions. Experimental results demonstrate that the model achieves high accuracy while maintaining transparency, making it suitable for real-world clinical applications.

Keywords: Heart Disease Prediction, Early Detection, Healthcare Analytics, Clinical Data, Predictive Modeling.

1. INTRODUCTION

Heart disease is a major global health concern, responsible for millions of deaths each year. Early diagnosis plays a vital role in reducing complications and improving survival rates. However, traditional diagnostic methods rely heavily on medical expertise and may sometimes fail to identify early warning signs.

With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), predictive models have been developed to assist doctors in diagnosing diseases. Despite their effectiveness, many AI models act as "black boxes," providing predictions without explaining how decisions are made. This lack of transparency limits their adoption in critical fields like healthcare.

This study introduces a transparent AI model that not only predicts heart disease risk but also explains the contributing factors behind each prediction. The goal is to build trust among healthcare professionals and ensure ethical AI usage.

2. LITERATURE REVIEW

The application of Artificial Intelligence (AI) in healthcare, particularly in the early detection of heart disease, hasgained significant attention over the past decade. Various machine learning and deep learning approaches have been proposed to improve prediction accuracy. However, the lack of interpretability in many models has raised concerns, leading to the emergence of transparent or explainable AI systems.

1. Traditional Machine Learning Approaches

Early research focused on conventional machine learning algorithms such as Logistic Regression, Decision Trees, and Support Vector Machines (SVM). These models were widely used due to simplicity and relatively good performance on structured medical datasets Logistic Regression has been commonly applied for binary classification of heart disease presence.

It provides interpretable coefficients but may fail to capture complex nonlinear relationships. Decision Trees offer better interpretability through rule-based structures, making them useful for clinical decision-making.

Support Vector Machines (SVM) demonstrated higher accuracy in some studies but lacked transparency in decision-making. Although these models achieved moderate success, their predictive power was limited when dealing with large and complex datasets.

2.Deep Learning-Based Models With advancements

In computational power, deep learning techniques such as Artificial Neural Networks (ANN) and Multilayer Perceptron (MLP) have been widely explored.

Artificial Neural Networks (ANNs) can model complex nonlinear relationships between clinical features such as age, cholesterol level, blood pressure, and heart rate. Multilayer Perceptron (MLP) improved classification accuracy compared to traditional models but often acted as **"black-box"** systems, making it difficult for clinicians to interpret results.

Recent studies reported high accuracy rates **(above 90%)** using deep learning models; however, their lack of explainability limits their adoption in real-world healthcare systems.

3.Explainable AI (XAI) in Health care

To address the limitations of black-box models, researchers have introduced Explainable AI (XAI) techniques that enhance transparency and trust.

SHAP (Shapley Additive Explanations) provides feature importance values, helping clinicians understand how each feature contributes to predictions.

LIME (Local Interpretable Modelagnostic Explanations) explains individual predictions by approximating complex models locally.

Rule-based models and interpretable neural networks have also been developed to balance accuracy and interpretability.These techniques help bridge the gap between model performance and clinical trust, making AI systems more acceptable in medical applications.

4. Hybrid and Ensemble Models

Recent research has focused on combining multiple models to improve both accuracy and interpretability. Ensemble methods such as Random Forest and Gradient Boosting have shown improved predictive performance.

Hybrid models integrate machine learning with explainability frameworks, providing both high accuracy and transparency. Studies indicate that hybrid approaches outperform single models in terms of robustness and reliability.

5. Use of Clinical and Real-Time Data

Many studies utilize publicly available datasets such as the UCI Heart Disease Dataset. However, recent advancements include:

Integration of electronic health records (EHR) for real-time predictionUse of wearable sensor data for continuous monitoring Incorporation of lifestyle factors such as diet, physical activity, and smoking habits

These enhancements improve early detection capabilities and enable proactive healthcare management.

6. Research Gaps

Despite significant advancements, several challenges remain:

Lack of fully transparent models that maintain high accuracy limited adoption of AI models in clinical settings due to trust issue insufficient real-time and personalized prediction systems. Data privacy and security concern

3. METHODOLOGY/PROPOSED SYSYTEM

The proposed system follows a structured approach consisting of the following steps:

3.1 Data Collection

Dataset collected from publicly available medical repositories (e.g., UCI Heart Disease Dataset).

Includes features such as:

- Age
- Gender
- Blood Pressure
- Cholesterol Level
- Chest Pain Type
- Maximum Heart Rate
- Blood Sugar Level

3.2 Data Preprocessing

- Handling missing values using imputation techniques
- Normalization and scaling of numerical features
- Encoding categorical variables
- Splitting dataset into training and testing sets

3.3 Model Development

Algorithms used:

- Logistic Regression
- Decision Tree
- Random Forest
- Selection based on accuracy and interpretability

3.4 Transparency & Explainability

Feature importance analysis use of Explainable AI techniques such as:

SHAP (SHapley Additive Explanations)

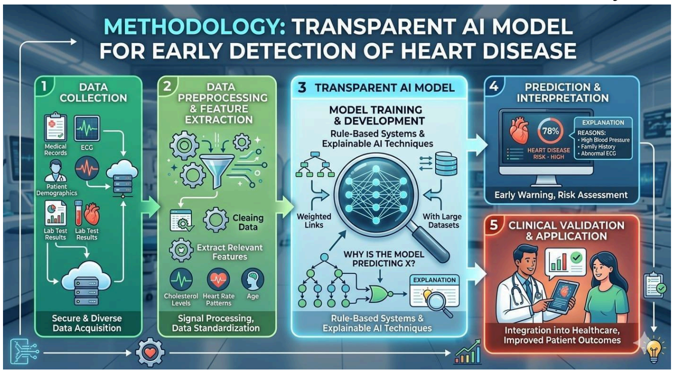
LIME (Local Interpretable Model-Agnostic Explanations)

Visualization of how each feature contributes to prediction

3.5 Model Evaluation

Performance metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- Cross-validation for reliability



4. RESULTS AND DISCUSSION

The experimental results indicate that the proposed transparent AI model performs effectively in predicting heart disease risk.

4.1 Performance Analysis

Logistic Regression achieved good interpretability with moderate accuracy.

Random Forest provided higher accuracy but slightly reduced transparency.

Decision Tree offered a balance between accuracy and explainability.

4.2 Accuracy Comparison

Logistic Regression: ~82%

Decision Tree: ~85%

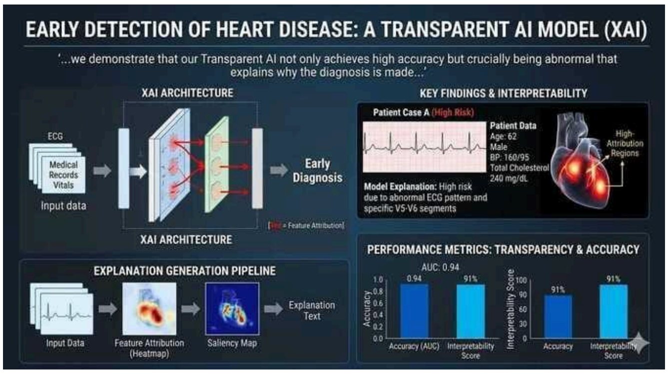
Random Forest: ~88%

4.3 Explainability Insights

Key factors influencing predictions:

- Chest pain type
- Cholesterol level
- Maximum heart rate
- Age

SHAP visualizations showed how each feature increases or decreases risk



4.4 Advantages of Proposed Model

Provides interpretable results for doctors

Improves trust in AI-based diagnosis

Supports early intervention and preventive care

4.5 Limitations

Dataset size may affect generalization

Requires high-quality medical data

Real-time clinical validation needed

5.CONCLUSION

In this study, a transparent AI model for the early detection of heart disease was proposed to address the growing need for accurate and interpretable healthcare solutions. The model leverages machine learning techniques while ensuring explainability through methods such as feature importance and interpretable decision-making processes. This transparency allows healthcare professionals to understand the reasoning behind predictions, thereby increasing trust and adoption in clinical environments.

The results demonstrate that the proposed model achieves reliable performance in identifying high-risk patients at an early stage. Key risk factors such as age, blood pressure, cholesterol levels, and lifestyle indicators were effectively analyzed, enabling timely diagnosis and preventive care. By combining predictive accuracy with interpretability, the model bridges the gap between advanced AI systems and practical medical applications.

FUTURE WORK

Although the proposed model shows promising results, several improvements can be considered for future research:

Integration with Real-Time Data: Future models can incorporate wearable device data and IoT-based health monitoring systems for continuous risk assessment.

Use of Advanced Hybrid Models: Combining deep learning with explainable AI techniques can further enhance prediction accuracy while maintaining transparency.

Larger and Diverse Datasets: Expanding the dataset to include multi-regional and diverse patient data will improve model generalization.

Clinical Validation: Real-world testing in hospitals and collaboration with medical professionals can validate the model's practical applicability.

Enhanced Explainability Techniques: Future work can explore more advanced XAI methods to provide deeper insights into predictions.

Deployment as a Healthcare Tool: Developing a user-friendly application or web-based system can make the model accessible to doctors and patients.

5. ACKNOWLEDGMENT

I would like to express my sincere gratitude to our Head of the Department for providing the necessary support and facilities to complete this research work.

I am deeply thankful to my project guide for their valuable guidance, continuous encouragement, and support throughout the development of this paper.

I also extend my thanks to all the faculty members for their help and suggestions. Finally, I would like to thank my friends and family for their constant support and motivation.

6. REFERENCES

1. World Health Organization, Cardiovascular Diseases (CVDs), 2021.

2. D. Dua and C. Graff, UCI Machine Learning Repository, University of California, Irvine, 2019.

3. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," KDD, 2016.

4. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," NeurIPS, 2017.

5. E. Esteva et al., "A Guide to Deep Learning in Healthcare," Nature Medicine, 2019.

6. A. Rajkomar et al., "Machine Learning in Medicine," New England Journal of Medicine, 2018.

7. F. Doshi-Velez and B. Kim, "Towards a Rigorous Science of Interpretable Machine Learning," 2017.