



EXPLAINABLE AI MODELS FOR CARDIOVASCULAR RISK ASSESSMENT

Vaibhav Jeevanrao Bhandari

*Department of Computer Science and Engineering
G H Rasoni University, Amravati, Maharashtra.*

Prashant K. Adakane

*Department Computer Science and Engineering
G H Rasoni University, Amravati, Maharashtra, India.*

1. Abstract

Cardiovascular diseases (CVDs) are among the leading causes of death worldwide. Early prediction and prevention are essential to reduce mortality rates. Artificial Intelligence (AI) and Machine Learning (ML) techniques have been widely used to predict cardiovascular risk using patient health data. However, many AI models act as “black boxes,” making their predictions difficult to understand for medical professionals. Explainable Artificial Intelligence (XAI) provides transparency by explaining how AI models make predictions. This paper presents an explainable AI framework for cardiovascular risk assessment using machine learning models and interpretation techniques such as SHAP and LIME.

Keywords— *Explainable Artificial Intelligence (XAI), Cardiovascular Disease (CVD) Prediction, Machine Learning, SHAP and LIME, Healthcare Decision Support Systems.*

2. Introduction

Cardiovascular diseases include heart attack, stroke, and coronary artery disease. These conditions affect millions of people every year. Early risk prediction can help doctors take preventive measures and improve patient care. Machine learning algorithms can analyze medical datasets and detect hidden patterns related to cardiovascular diseases. However, healthcare professionals require explanations for AI predictions before adopting them in real clinical environments. Explainable Artificial Intelligence (XAI) helps make machine learning models transparent and interpretable, enabling doctors to understand the factors influencing predictions.

3. Problem Statement

Traditional machine learning models provide high accuracy but lack interpretability. Doctors cannot easily understand why the model predicts a patient as high-risk or low-risk. This lack of transparency limits the adoption of AI systems in healthcare. Therefore, there is a need for explainable AI models that

provide both accurate predictions and clear explanations for cardiovascular risk assessment.

4. Objectives

1) Develop machine learning models for cardiovascular risk prediction

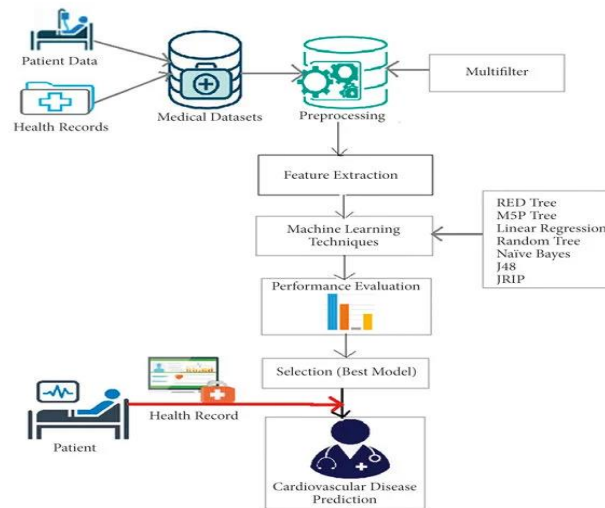


Fig1. Develop machine learning models for cardiovascular risk prediction

1. Patient Data & Health Records

- Data such as age, blood pressure, cholesterol, diabetes history, and lifestyle information are collected from patients.

2. Medical Dataset Creation

- All patient information is stored together to form a medical dataset used for analysis.

3. Preprocessing

- The dataset is cleaned by removing errors, missing values, and irrelevant information.
- Multifilter techniques may be used to improve data quality.

4. Feature Extraction

- Important health features (e.g., cholesterol level, heart rate, BMI) are selected from the dataset for model training.

5. Machine Learning Techniques

- Different algorithms are applied, such as:
 - Decision Tree (RED Tree, M5P Tree)
 - Linear Regression
 - Random Tree
 - Naïve Bayes
 - J48
 - JRIP

6. Performance Evaluation

- Each model is evaluated using performance metrics like accuracy, precision, recall, and F1-score.

7. Best Model Selection

- The model with the highest performance is selected as the final predictive model.

8. Prediction

- The selected model analyzes a new patient's health record and predicts the risk of cardiovascular disease.

2) Apply explainable AI techniques for model interpretation

Explainable Artificial Intelligence (XAI) techniques are used to understand and interpret how machine learning models make predictions. In healthcare applications such as cardiovascular disease (CVD) risk prediction, XAI helps doctors and researchers identify which factors influence the model's decision. This improves transparency, reliability, and trust in AI systems.

1. Purpose of XAI

- To explain the predictions made by complex AI models.
- To identify important features affecting the prediction.
- To increase transparency and trust in healthcare AI systems.
- To help medical professionals validate AI decisions.

2. Common XAI Techniques

1. SHAP (Shapley Additive Explanations)

- Based on game theory.
- Shows the contribution of each feature to the prediction.
- Helps identify important factors such as age, blood pressure, cholesterol, smoking, etc.

2. LIME (Local Interpretable Model-Agnostic Explanations)

- Explains individual predictions of a machine learning model.
- Creates a simple interpretable model around a specific prediction.
- Helps understand why the model predicted high or low CVD risk for a particular patient.

3. Feature Importance

- Measures how much each feature contributes to the model's predictions.
- Helps rank medical factors like blood pressure, glucose, weight, and cholesterol.

4. Partial Dependence Plots (PDP)

- Shows the relationship between a feature and the predicted outcome.
- Helps visualize how a variable affects CVD risk.

5. Decision Tree Visualization

- Provides a clear graphical representation of decision paths in tree-based models.

3. Application in CVD Risk Prediction

When XAI is applied to a cardiovascular prediction model:

- The model predicts whether a patient has high or low heart disease risk.
- XAI methods explain which medical factors influenced the prediction.
- Doctors can verify if the AI decision is medically reasonable.

4. Benefits of Using XAI

- Improves model transparency
- Builds trust in AI healthcare systems
- Helps doctors understand AI predictions
- Supports better clinical decision-making

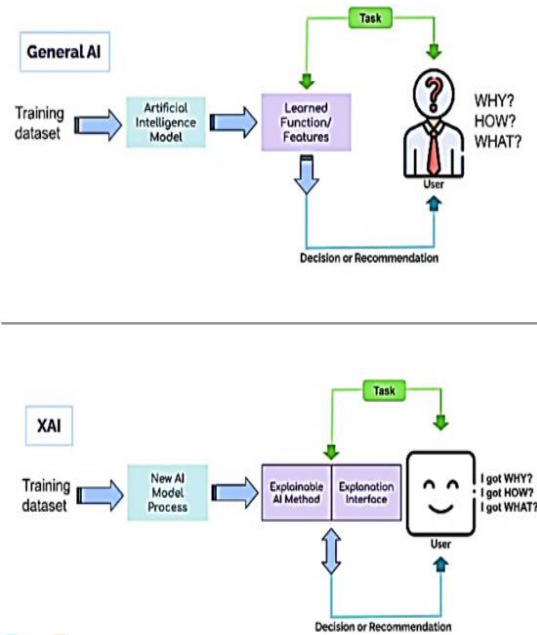


Fig.2 Explainable AI techniques for model interpretation

. General AI (Top Diagram)

- A training dataset is used to train an AI model.
- The model learns features and patterns from the data.
- It produces a decision or recommendation for a task.
- However, the user cannot understand the reason behind the decision.
- Users often ask “Why?”, “How?”, and “What?”, but the model does not clearly explain it.

. Explainable AI (XAI) (Bottom Diagram)

- The training dataset is used in a new AI model process with explainability methods.
- Techniques from **Explainable Artificial Intelligence provide explanations of the model’s predictions.
- An explanation interface shows how the decision was made.
- The user can clearly understand Why the prediction happened, How it was calculated, and What factors influenced it.

3) Identify key medical factors affecting cardiovascular risk

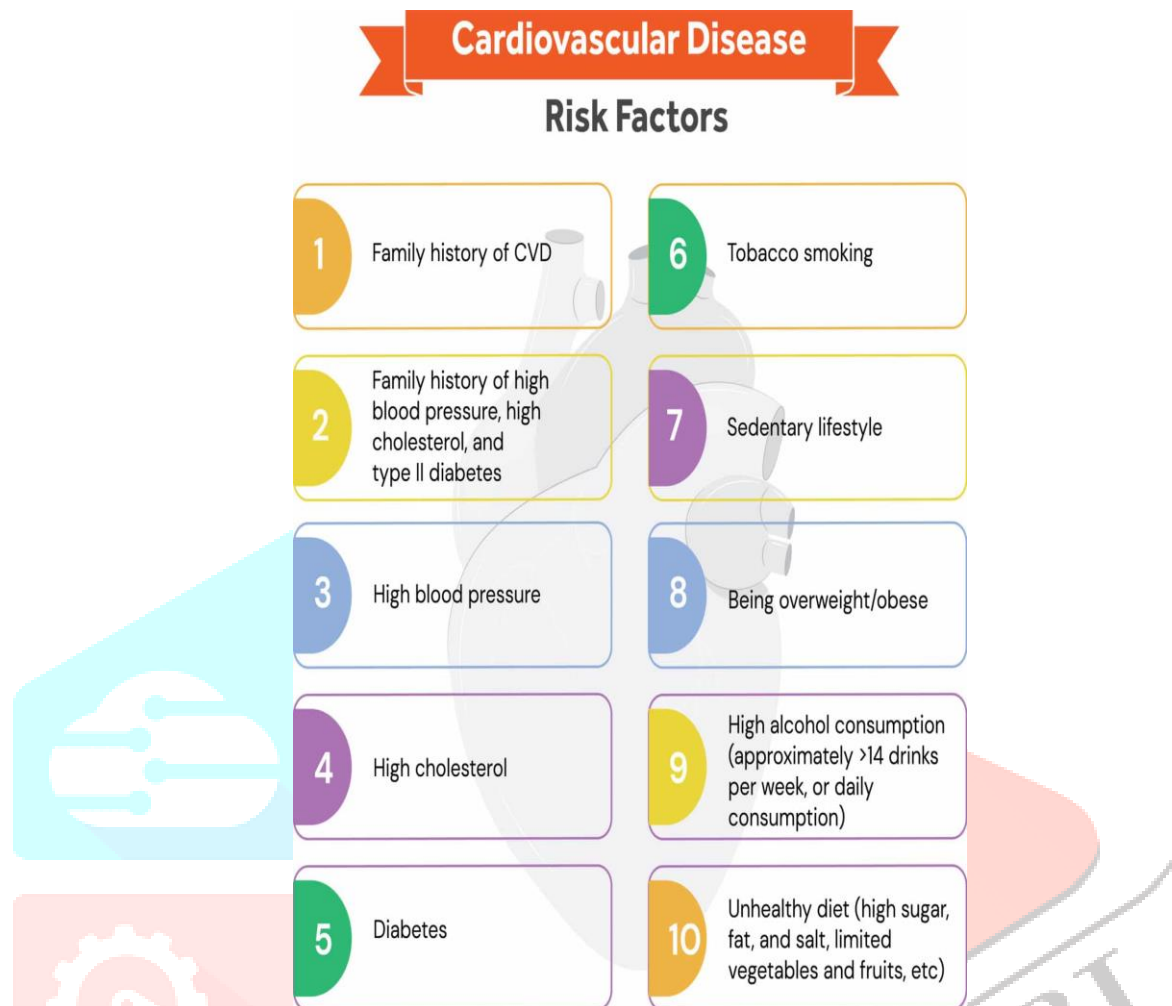


Fig 3. Identify key medical factors affecting cardiovascular risk

- **Family history of CVD** – If heart disease runs in the family, the risk increases.
- **Family history of BP, cholesterol, or diabetes** – Genetic factors may lead to these conditions.
- **High blood pressure** – Increases strain on the heart and arteries.
- **High cholesterol** – Causes fatty deposits in blood vessels, blocking blood flow.
- **Diabetes** – High blood sugar damages blood vessels and the heart.
- **Tobacco smoking** – Damages blood vessels and increases heart disease risk.
- **Sedentary lifestyle** – Lack of physical activity weakens heart health.
- **Overweight or obesity** – Excess body weight increases heart workload.
- **High alcohol consumption** – Excess drinking raises blood pressure and heart risk.
- **Unhealthy diet** – High sugar, fat, and salt intake can lead to heart disease.

4. Improve transparency and trust in healthcare AI systems

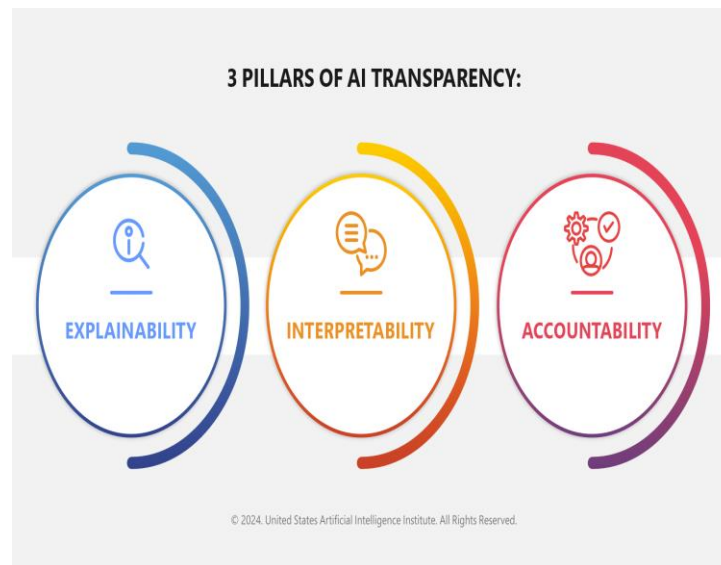


Fig 4. Improve transparency and trust in healthcare AI systems

1. Explainability

- Explainability means the AI system can clearly explain why it made a particular prediction or decision.
- It shows which features or data influenced the result.
- Example: In heart disease prediction, the model explains that high blood pressure and cholesterol increased the risk.

2. Interpretability

- Interpretability means humans can easily understand how the AI model works.
- The relationship between input data and output results is clear and understandable.
- Simple models like decision trees or linear regression are often more interpretable

3. Accountability

- Accountability means someone is responsible for the AI system's decisions and outcomes.
- Organizations must ensure the AI model is fair, safe, and used ethically.
- If errors occur, there must be clear responsibility and correction.

5. Cardiovascular Risk Factors

Important factors include Age, Gender, Blood Pressure, Cholesterol Level, Blood Sugar Level, Smoking Habit, Alcohol Consumption, Physical

Activity, and Body Mass Index (BMI). These parameters are used as input features in machine learning models.

6. Dataset

Public healthcare datasets are used such as Kaggle Cardiovascular Disease Dataset and UCI Heart Disease Dataset. Attributes include age, gender, height, weight, blood pressure, cholesterol, glucose level, and smoking status.

7. Methodology

Steps involved: Data Collection → Data Preprocessing → Feature Selection → Model Training → Risk Prediction → Explainable AI Analysis → Result Interpretation

8. Machine Learning Models

Logistic Regression – Binary classification. Decision Tree – Rule-based tree model.

Random Forest – Ensemble of multiple trees.

Support Vector Machine – High dimensional classification.

9. Explainable AI Techniques

SHAP explains feature importance such as age, blood pressure, and cholesterol influence. LIME explains individual predictions of machine learning models.

10. Results and Discussion

Random Forest and SVM often achieve higher accuracy. Explainable AI highlights major risk factors like age, blood pressure, cholesterol, and glucose levels, helping doctors understand model decisions.

11. Applications

Hospital decision support systems, early disease detection, preventive healthcare programs, personalized medicine, and remote health monitoring.

12. Advantages

Early detection of heart disease, transparent AI decision-making, improved doctor trust, and better healthcare planning.

13. Future Scope

Future work includes deep learning explainability, wearable health device integration, real-time monitoring, and personalized healthcare systems.

14. Conclusion

Explainable AI improves transparency and trust in machine learning models used in healthcare. Combining prediction models with explanation techniques supports early diagnosis and prevention of cardiovascular diseases.

15. References

1. Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *NeurIPS*.
2. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. *ACM KDD*.
3. Benjamin, E. J., et al. (2019). Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association.
4. Shickel, B., et al. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*.
5. World Health Organization (WHO). (2024). Cardiovascular diseases (CVDs) fact sheet. Retrieved from [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
6. Benjamin, E. J., Muntner, P., Alonso, A., et al. (2019). Heart Disease and Stroke Statistics—2019 Update: A Report from the American Heart Association.
7. *Circulation*, 139(10), e56e528. <https://doi.org/10.1161/CIR.0000000000000659>
8. UCI Machine Learning Repository. (2023). Heart Disease Data Set. University of California, Irvine. Retrieved from <https://archive.ics.uci.edu/ml/datasets/heart+Disease>
9. Dey, D., Slomka, P. J., Leeson, P., Comaniciu, D., Shrestha, S., & Sengupta, P. P. (2019). Artificial Intelligence in Cardiovascular Imaging:
10. JACC State-of-the-Art Review. *Journal of the American College of Cardiology*, 73(11), 1317–1335. <https://doi.org/10.1016/j.jacc.2018.12.054>
11. Krittanawong, C., Johnson, K. W., Rosenson, R. S., et al. (2021). Deep learning for cardiovascular medicine: A practical primer. *European Heart Journal*, 42(12), 11701183. <https://doi.org/10.1093/eurheartj/ehaa762>

12. Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*, 29(2), 102–12.

