

Heart Disease Prediction in Explainable AI-Based Clinical Decision Support System based on XGBoost and SHAP

Ravi Ranjan Singh¹, Jatin Kumar², Bijoyini Bagchi²,
Neha Gupta², Aditya Singh²

Computer Science and Engineering, Institute of Engineering and Management, Kolkata

Abstract: Predicting heart disease using clinical data can provide valuable information for early risk identification, and the difficulty of many machine learning models to interpret is a challenge in healthcare applications. This research proposes an interpretable machine learning model to compare Logistic Regression, Decision Tree, Random Forest, and XGBoost models for predicting heart disease. The clinical data are preprocessed and then split into a training set and a test set, then the model is trained, and finally compared with the accuracy, precision, recall and F1 score. The optimized XGBoost model outperformed the others, making the accuracy (90.76). Subsequently, the clinical features that lead to the prediction of the model were identified using the SHapley Additive exPlanations (SHAP) method. The design proposed here ensures predictive performance and allows for interpretability of the results and it can be used as a potential basis for clinical decision-support applications.

Keywords: Heart Disease Prediction , Machine Learning , Explainable AI , XGBoost , SHAP , Clinical Decision Support.

1 Introduction

Heart disease is still a significant health problem and early risk detection is crucial for early intervention. The clinical data have been extensively used for machine learning techniques like Logistic Regression, Decision Tree, Random Forest and XGBoost. But complex models can be uninterpretable, which means that they cannot be accepted in healthcare. To overcome this limitation, approaches such as LIME or SHAP were created which are able to explain the impact of individual features. The goal of this study is to optimize the performance of the classifiers while maintaining interpretability while accurately predicting the presence of heart disease, this study evaluates four classifiers and combines optimized XGBoost with SHAP.

2 Literature Review

The field of machine learning, clinical data has been employed in numerous applications, such as predicting heart diseases. Detrano et al. created a diagnostic model based on probabilities that contributed to the creation of the well known Heart Disease database of the UCI [1,9]. In classical supervised learning techniques, Kotsiantis studied them and explained their appropriateness for the medical classification problem as well as the fact that they are unable to handle complex nonlinear relationships [6]. Chen and Guestrin proposed an efficient gradient-boosting algorithm, XGBoost, which is good at prediction on structured data [3]. Advanced machine learning models, however, are still a big challenge in healthcare, in being interpretable. Ribeiro et al. [5] proposed the concept of LIME for explaining individual predictions and Lundberg and Lee [4] proposed the concept of SHAP to provide the contribution of the features to the output of the model. However, there is a challenge for research in the development of good predictive models that are clinically relevant. To tackle this, this paper tests a variety of classifiers and uses SHAP for the optimized XGBoost.

3 Methodology

3.1 System Overview

The proposed pipeline comprises six steps: Data collection, Data Preprocessing and Feature Transformation, splitting the data into (80 percentage) train and (20 percentage) test data, training the model, evaluating the model and explainability analysis using SHAP.

3.2 Dataset and Preprocessing

Other laboratory values and some of the clinical features from each patient are available in these data, for example, whether the person has chest pain, their blood pressure when they're not working out, whether they've got cholesterol, and more. Duplicate records were removed, and only the relevant features were retained and StandardScaler then applied. The dataset was then divided into two sets: 80% for training and 20% as a test set.

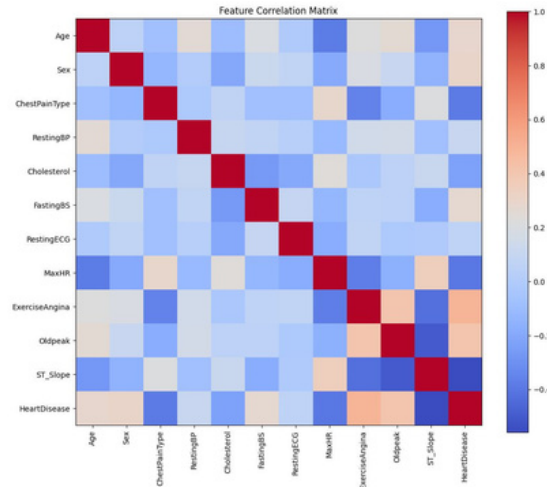


Figure 1: Feature correlation matrix, showing how clinical attributes relate to one another and to the heart-disease target.

3.3 Models and Explainability

The classifiers that were trained are: Logistic Regression, Decision Tree, XGBoost and Random Forest. Hyperparameter optimized XGBoost was used for the main model. All the features were ranked based on their importance using SHAP and individual predictions and overall feature contributions were explained.

3.4 Evaluation Metrics

Accuracy, precision, recall and F1 score were used to assess the performance of the model. Confusion matrix was also worked out for analysing true and false classifications.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad \text{Precision} = \frac{TP}{TP + FP}$$
$$\text{Recall} = \frac{TP}{TP + FN}, \quad F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4 Experimental Setup

The experiments have been carried out in python, using Scikit learn and XG boost, in google Colab. Preprocessed data then was split into input features and the target (heart-disease) class, which is binary. RandomizedSearchCV for XGBoost optimization and comparing the classification performances based on the accuracy, precision, recall, F1-score with other classifiers. In addition, a confusion matrix and SHAP feature importance plot was created.

5 Results and Discussion

5.1 ModelComparison

The accuracy, precision, recall and F1 score is demonstrated for the four classifiers in Table 1 below. The XGBoost model was the one that is most optimized and it was seen to perform best with accuracy of (90.76) while the RF model was the second best model. The results showed that the tree-based ensemble models performed better than Logistic Regression and Decision Tree on the given clinical data.

Table 1: Performance comparison of the machine learning models (%).

Model	Accuracy	Precision	Recall	F1Score
LogisticRegression	86.96	83.00	87.00	85.00
DecisionTree	78.80	81.00	84.00	82.00
RandomForest	89.67	89.52	92.16	90.82
XGBoost	90.76	91.26	92.16	91.71

5.2 Confusion Matrix and SHAP Analysis

The confusion matrix for the optimized XGBoost model shows that in most cases, the model had accurately determined if a row had data for a disease or not. SHAP was used to explore how individual clinical metrics relate to the model's predictions. For the analysis of individual clinical features in the model output, SHAP analysis was used. The analysis provides an understandable view of what conditions predict heart disease, and can be combined with the quantitative performance metrics.

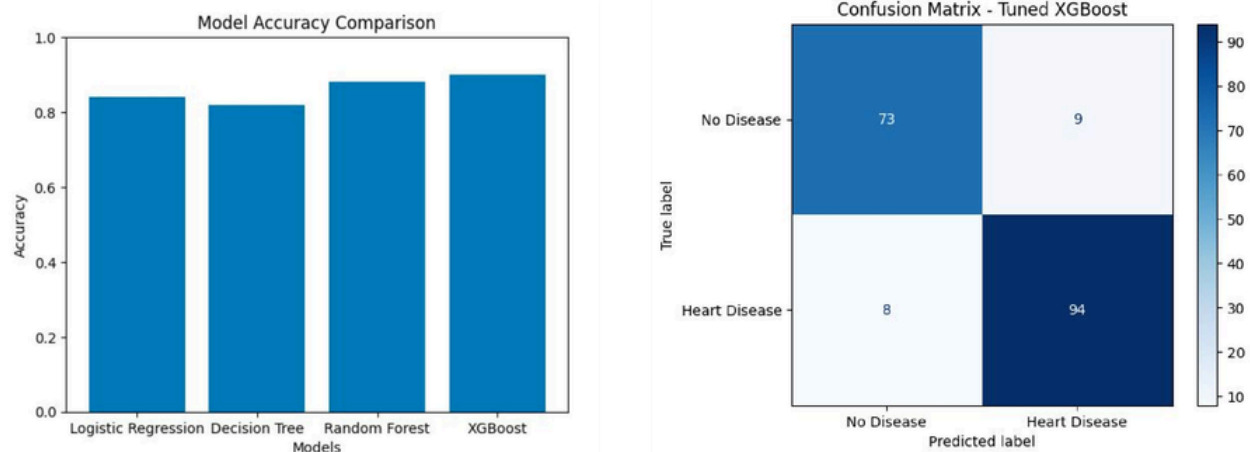


Figure 2: Accuracy comparison across the four classifiers (left) and confusion matrix of the optimized XGBoost classifier (right).

5.3 Discussion

Overall, the results show that the optimized XGBoost classifier is a well-performing machine learning model that enables the effective prediction of the issue, while the SHAP model shows the improved interpretability of the model's predictions. The two are complementary in their development of clinically useful machine learning systems for clinical decision support in healthcare.

6 Conclusion

In this research, machine learning methods were applied for predicting heart disease using clinical parameters. The numeric order of classifiers based on their overall performance was optimized- XGBoost classifiers (90.76%), Random Forest Classifier (89.67%), Logistic Regression Classifier (86.96%) and as the least best was the Decision Tree Classifier with (78.80%). Interpretability of the model was also shown with SHAP giving interpretable view, showing which feature of the model was responsible for each prediction. The results show how well the proposed approach of predictive machine learning and explainable AI can help improve the development of more explainable and reliable clinical decision-support systems.

7 Future Work

Future research will concentrate on enhancing robustness and generalization with larger and more varied clinical datasets that will be used to validate the framework. Further study and comparison might be conducted using other machine learning methods and deep learning methods. Some other ways of explaining could be explored to provide similar explanations. The system might be further extended in the future to a real-time clinical decision-support system to help health care professionals identify risk at an early stage.

References

- [1] Detrano, R., et al.: International application of a new probability algorithm for the diagnosis of coronary artery disease. *Am. J. Cardiology* 64(5), 304–310 (1989)
- [2] Breiman, L.: Random forests. *Machine Learning* 45(1), 5–32 (2001)
- [3] Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD*, 785–794 (2016)
- [4] Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. *NeurIPS*, 4765–4774 (2017)
- [5] Ribeiro, M.T., Singh, S., Guestrin, C.: “Why should I trust you?” *Proc. 22nd ACM SIGKDD*, 1135–1144 (2016)
- [6] Kotsiantis, S.B.: Supervised machine learning: A review. *Informatica* 31(3), 249–268 (2007)
- [7] Cortes, C., Vapnik, V.: Support-vector networks. *Machine Learning* 20(3), 273–297 (1995)
- [8] Quinlan, J.R.: Induction of decision trees. *Machine Learning* 1(1), 81–106 (1986)
- [9] Dua, D., Graff, C.: UCI Machine Learning Repository. Univ. of California, Irvine (2019)
- [10] Lundberg, S.M., Erion, G., Lee, S.I.: Consistent individualized feature attribution for tree ensembles. *arXiv:1802.03888* (2018)
- [11] Ahmad, M.A., Eckert, C., Teredesai, A.: Interpretable machine learning in healthcare. *Proc. 2018 ACM BCB*, 559–560 (2018)
- [12] Deo, R.C.: Machine learning in medicine. *Circulation* 132(20), 1920–1930 (2015)
- [13] Rajkomar, A., Dean, J., Kohane, I.: Machine learning in medicine. *N. Engl. J. Med.* 380(14), 1347–1358 (2019)