HEART ATTACK PREDICTION

Challenging Assignments and Mini Projects (CHAMP)

Submitted By
Siddharth Ajay Gor 21BAI1670

Abstract

Heart disease is still one of the leading causes of mortality worldwide, and early prediction of heart attacks is crucial for effective intervention and prevention. This paper proposes a predictive model using machine learning techniques in general, and, in particular, the SHAP method for transparent and interpretable insight into feature importance. Our approach integrates diverse data of the patient from demographic, medical history, and lifestyle for a comprehensive predictive model. The model can thus predict well the likelihood of a heart attack occurrence by using some feature engineering and ML algorithms, such as decision trees or gradient boosting. We further explain the contribution of each feature in the prediction using SHAP, which makes it even more interpretable and easy to derive actionable insights for healthcare practitioners. Real-world experimental results on the proposed methodology have shown efficacy and illustrated potential to help healthcare providers identify and implement, early on, intervention strategies for people at risk of heart attacks.

Keywords: *Heart Disease, Mortality, Early Prediction, Heart Attack, Machine Learning (ML), SHAP (SHapley Additive exPlanations), Feature Importance, Predictive Model, Patient Data, Demographics, Medical History*

Contents

1 Introduction

Equal rights for all and well-being of the populace are the underpinning principles of a just and fair society. It relates to adequate housing, an enabling environment that is healthy, and the essence of major healthcare services. By assuring these rights, the state provides an assurance of making all its citizens experience the highest standard of healthcare service. True, healthcare is essential in safeguarding the well-being of the people, pursuant to what the constitution provides; quality services provided must be there. The prominent leaps in technology humanity has taken have contributed greatly to the practices in healthcare and to improving lives. Pioneering innovations are changing the ways through which healthcare professionals diagnose and treat diseases and ultimately transforming patient care outcomes. A good example is the application of Magnetic Resonance Imaging (MRI) machines to very accurately identify various medical conditions. By blending science that processes data together with electronic devices, health practitioners have at their disposal a tool to be able to diagnose accurately and attempt to cure health worries in their patients.

Technology integration shall bring forth unprecedented opportunities that allow healthcare to delve into early detection, accurate diagnosis, and personalized planning of treatment. The use of advanced imaging techniques helps doctors identify subtle abnormalities that would otherwise be nearly impossible to detect, such as those visible in the MRI scan. Early detection allows timely interventions, which might save lives or improve long-term health outcomes. Moreover, the seamless convergence between the science of data handling and the electronic gadgets has resulted in a much more efficient and seamless healthcare system. It was established that electronic information storage could provide the twin capability of security and access to patient's information electronically, which when integrated facilitated systematic and coordinated care in different settings for healthcare. One issue associated with these types of machine learning models is that not all the elements associated with models have clear interpretability. The key focus at the time of creation of the model was to ensure and enhance the accuracy of predictions. However, in the medical field, interpretation remains an important aspect if at all someone is to understand and use the models appropriately. For correct diagnosis of disease in the patient, there must be an interpretation of the model in question. This should come with paying attention to detail. Effort should be directed toward explaining the predictive model used for diseases. In this work, we will continue to develop a high-accuracy prediction model for heart disease that is interpretable. This study will use SHAP to find out the factors causing a positive patient diagnosis.

1.1 Problem

The challenge of ensuring equal access to quality healthcare is critical in creating a just and equitable society. Despite advancements in medical technology, significant disparities remain in access to essential services, particularly for life-threatening conditions like heart disease, which is a leading cause of death globally. Early detection and accurate diagnosis are crucial, yet many predictive models, while advanced, lack transparency and interpretability, limiting their effectiveness for healthcare professionals. This complexity makes it difficult for providers to make informed decisions that could save lives. Additionally, the integration of technology into healthcare systems raises concerns about data security, accessibility, and equitable distribution of resources. The problem lies not only in developing sophisticated medical technologies but also in ensuring these tools are accessible, understandable, and effectively used to improve health outcomes for all individuals, regardless of their socio-economic status or location.

1.2 Scope

The project will involve the development of an interpretable machine learning model for predicting heart attacks that could be integrated into healthcare systems to help in early diagnosis and intervention. More precisely, the project addresses the dual needs identified above: those of predictive accuracy and model interpretability. These ensure that the healthcare provider can trust the prediction and understand the reasoning behind it. It involves demographic information about the patient, past medical history, and lifestyle factors. This model takes a broad array of inputs in order to provide a detailed diagnosis of heart attack risk. The problem at hand is complex in nature and considers high-dimensional datasets, which involve decision trees and ensemble methods like gradient boosting. Additionally, the SHAP method was used to give interpretability to the model's predictions and results, enabling healthcare professionals to understand the most influential factors in assessing risk.

The domain of the scope also covers a model being integrated into healthcare, focusing on how medical practitioners may find it user-friendly. This model may be helpful in guiding clinical decisions and preventive strategies since it will give a clear breakdown of all the factors involved in each prediction. Besides, the model is transparent and hence can be applied in wide-ranging healthcare settings: from the smallest clinics to big hospitals, without the need for active training among medical personnel.

Beyond the technical, this work further discusses the wide raft of ramifications implicated in the implementation of machine learning into healthcare. Issues to do with data security, patient privacy, and equity distribution of resources are paramount in this regard while implementing such models. The focus of this project is, therefore, on working out the effective model of the three commanding aspects: prediction, ethics, and accessibility.

2 Related Works

2.1 Improving Prediction Accuracy Using Advanced Machine Learning Techniques

Gao et al. [2] attempt to address the limitation of prediction accuracy found in Rairikar et al. [3] through proposing more advanced models of machine learning. The key highlights of the authors are on Support Vector Machines and Random Forest classifiers for heart disease prediction. This work investigates how model tuning and parameter optimization can result in significantly enhanced performance of earlier, simple models such as Decision Trees and Naive Bayes. Another contribution of Gao et al. is that they showed model hyperparameters and feature selection strategies could be optimized to improve the predictive power of machine learning models. The authors were able to show that even small changes in the parameters greatly enhance the accuracy with the introduction of cross-validation techniques and grid search methods. The Random Forest classifier, for example, has been very resistant to overfitting and produced very good results in the prediction task. Their study actually marks a turning point in the literature, passing from traditional data mining approaches to more specialized machine learning ones. However, Gao et al. [2] also point out that there are still some difficulties in the normalization of the input data. The quality and homogeneity of medical data are directly linked with the accuracy of the predictions; therefore, there is a need for one ontological framework in order to better structure the information related to diseases.

2.2 History

Heart disease is one of the major causes of death among people around the world. The rate of mortality from heart conditions can be decreased by early detection and diagnosis. With the emergence of machine learning and data mining, prediction or diagnosis of heart diseases has created new avenues. Various studies have been developed with the intention of creating algorithms that will analyze medical data efficiently in predicting heart diseases. This literature review traces the evolution of methodologies developed for the prediction of heart disease, from simple data mining techniques to complex machine learning approaches, ontological frameworks, and feature selection strategies.

2.3 Prediction of Heart Disease Using Data Mining Techniques

Rairikar et al. [3] present applications of data mining techniques in predicting heart diseases. The authors have used a number of algorithms, including Decision Trees, Naive Bayes, and K-Nearest Neighbors. These are benchmarked on a set of patients affected with heart disease. The objective is to find an efficient model for predicting heart disease with high accuracy. The paper shows that each algorithm has various strengths and weaknesses: for instance, Decision Trees are rather interpretable but might experience overfitting problems, while KNN is resistant to such a problem but rather computationally expensive. According to Rairikar et al. [3], methods of data mining give important insights; however, their predictive ability is far from being good enough for clinical purposes. They therefore conclude that further improvements are needed, such as advanced machine learning techniques. Their work forms the necessary basis for further studies that try to improve the accuracy by proposing more sophisticated methods.

2.4 Ontological Framework for Heart Disease Diagnosis

Scheuermann et al. [5] present a challenge in the standardization of data and semantic clarity by using an ontological treatment of disease and diagnosis. Their work focuses on the ontological development of a framework to offer a harmonized structure of representation of diseased data, necessary in enhancing the efficiency of machine learning models. This ontological treatment is particularly important in the healthcare field, where inconsistencies of a terminological nature and incomplete data can severely undermine predictive accuracy. Also, diseases such as heart disease require more than statistical and computational models for diagnosis; they require that those be captured through some rigorous and semantically consistent representation. The ontological framework presented in the paper allows a better classifying of medical data in a consistent way and with a link to clinical findings. Formal organization of data into ontology form will make the recognition of patterns by machine learning algorithms easier and hence hopefully lead to better prediction and diagnosis. It provides a sound structure for predictive models to work on, complementing recent advances in machine learning algorithms themselves, described in Gao et al. [2]. The healthcare data is growing into an increasingly complex entity with each passing day; hence, these ontological frameworks become all the more urgent. In this manner, Scheuermann et al. provide an important contribution to establishing links between clinical data to computational models for more reliable predictions.

2.5 Feature Selection and Classification Techniques

The success so far recorded in machine learning and ontological frameworks motivated Reddy et al. [4] to study the significant role of feature selection in enhancing the accuracy of heart disease prediction. Their investigation explores a number of classification techniques, including Decision Trees, SVM, and Neural Networks, which have been considered in their application coupled with feature selection methods. The authors also argue that the irrelevant or redundant features are one of the main reasons for the poor prediction accuracy of earlier studies. Feature selection methods, such as RFE and PCA, remarkably improve the performance of the model. Feature selection decreases the dimensionality of the data so that the adopted machine learning algorithms can focus their resources on the most important attributes, like age, cholesterol level, and blood pressure, critically associated with heart disease. Reddy et al. [4] illustrate the highest accuracy in prediction for models using SVM combined with RFE over the traditional classifiers. This paper further lends credence to the underlying reason why feature selection is a crucial ingredient in machine learning model improvements. The results showed the importance of balancing model complexity with a meaningful set of features so as to avoid overfitting and improve generalization.

2.6 Overall Review of Disease Diagnosis Based on Machine Learning

Ahsan et al. give an overview regarding the progress in ML-based disease diagnosis, where much focus is given to heart disease. Their study encapsulates the findings from different studies and presents a bird's-eye view of how machine learning could be adopted for accurate diagnosis. It provides a general overview of various machine learning algorithms, both unsupervised and supervised, applied to the diagnosis of heart diseases, diabetes, cancer, and other chronic illnesses. Ahsan et al. [1] located several open problems that have yet to be considered for machine learning in the health domain: interpretability, bias in training data, integration of heterogeneous sources. They have also indicated feature selection and ontological frameworks, to which correspond the works of Reddy et al. [4] and those of Scheuermann et al. [5], respectively. Ahsan et al. [1] conclude that, up to the time, adequate studies have been conducted to construct machine learning models for disease prediction; however, much more robust and interpretable models that could be trusted clinically are yet called for. They recommend including methods developed under the heading of XAI and continuing efforts on the application of ontologies for the development of machine learning models in healthcare that can provide both-accuracy and reliability.

2.7 Conclusion

The review of methodologies of heart disease prediction reveals an evident line of evolution: from rather simple data mining techniques to more sophisticated machine learning approaches. Rairikar et al. [3] gave the basics from the application of basic data mining methods, while successive research works such as Gao et al. [2] applied more powerful machine learning techniques, hugely improving predictive accuracy. Scheuermann et al. [5] then gave the ontological frameworks that can give a semantic clarity essential for the proper representation of data in machine learning models. Reddy et al. [4] further refined these methods to feature selection, while Ahsan et al. [1] then integrated such advances into the broad overview of the current state of machine learning-based disease diagnosis. Together, these papers show the step-by-step refinement of methodologies with an aim to further improve predictions of heart diseases. Every study adds to the knowledge base on crucial challenges concerning prediction accuracy, standardization of data, and selection of features. Future studies should be directed at integrating these machine learning techniques into a clinically useful model that would, therefore, be trusted by a health professional.

3 Design

The design of the predictive model includes several pivotal elements: data input, preprocessing, feature selection, model training, and interpretation using SHAP.

3.1 Data Input and Preprocessing

Data input consists of patient demographics, medical history, and lifestyle information. Preprocessing involves handling missing values, normalizing continuous variables, and encoding categorical variables.

3.2 Feature Selection

Feature selection is performed using methods like recursive elimination, aiming to identify the most relevant features for heart attack prediction.

In essence, this code snippet uses the SHAP library to explain the predictions of a Random Forest model by computing SHAP values for a sample of the test data. SHAP values help in understanding how each feature contributes to the model's predictions, providing insights into the model's behavior and potential areas for improvement.

Feature Importance: SHAP values measure the impact of each feature on the model's predictions. Healthcare professionals can analyze these SHAP values to identify which features (such as blood pressure, cholesterol levels, smoking status, etc.) are most influential in predicting the risk of a heart attack. Individual Prediction Explanation: SHAP provides explanations at the individual level, showing why the model assigned a certain risk score or probability of a heart attack for a given patient.

3.3 Model Training

The machine learning model is trained using decision trees and gradient boosting algorithms, ensuring the ability to capture complex patterns and interactions within the dataset. Cross-validation techniques are employed to enhance the model's generalization to unseen data, reducing the risk of overfitting. The SHAP library is then integrated to provide interpretability for the model's predictions. A `shap.Explainer` object is instantiated, initialized with the trained Random Forest model (`rf`) and the training dataset (`x_train`). This explainer computes SHAP values, which quantify the contribution of each feature to the model's predictions. SHAP values are calculated for a subset of the test dataset (`x_test`), illustrating how individual variables influence the prediction outcomes.

To optimize performance on large datasets, the parameter `check_additivity=False` is used, disabling additivity checks while retaining accurate explanations. This methodology ensures healthcare practitioners can interpret predictions effectively by understanding the impact of specific features on the risk assessment.

3.4 SHAP Integration

The SHAP method is integrated for model interpretability. SHAP values quantify the contribution of each feature to the prediction, offering transparency into why a particular patient is classified as high or low risk.

Table 2. F1-Score on each model.

Model	Positive (1)	Negative (0)
SVM	88%	85%
Random Forest	87%	82%
XGBoost	88%	85%
k-NN	86%	81%

Figure 1: Fig 1.0 F1 Score on each Model

4 Architecture

The architecture can be built upon a few pivotal elements: data input, preprocessing, feature selection, and model training and interpretation using SHAP.
Data input consists of patient demographics, the medical history of any particular patient, and lifestyle information. This data is preprocessed to handle possible missing values, normalize continuous variables, and encode categorical variables.
Feature selection is performed using methods such as recursive feature elimination, which identifies the most relevant features for heart attack prediction.
Once the data is prepared, the machine learning model is trained using a combination of decision trees and gradient boosting algorithms. These algorithms are selected for their ability to handle complex patterns and interactions between features. Additionally, cross-validation is employed to ensure the model generalizes effectively to new data.

4.1 UML Diagram

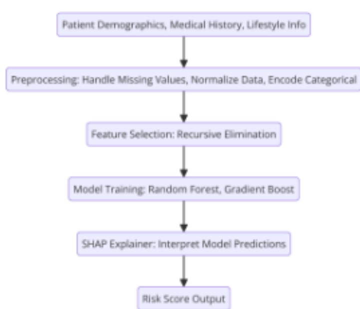

Figure 2: Interface Diagram 1

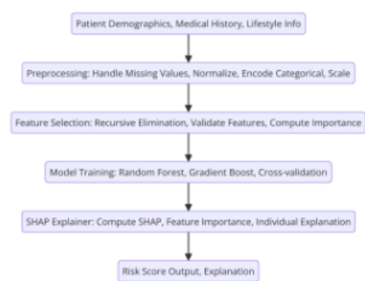

Figure 3: Interface Diagram 2

5 Observation

During the training and testing of the model, several important observations were made regarding factors contributing to heart attack risk. SHAP values identified the most valuable features, with age, cholesterol level, and family medical history emerging as the top predictors. Additionally, physical activeness and smoking status were influential features that significantly affected the model's predictions.
One of the most striking observations was the distinction in feature importance between high-risk and low-risk patients. For high-risk patients, medical history and age were major determining factors, whereas lifestyle factors had greater prominence for low-risk individuals. This suggests that individualized treatment plans must consider the specific medical and lifestyle inputs of each patient.

Moreover, SHAP revealed unexpected feature interactions, such as the dependency between blood pressure and cholesterol levels. These insights hold the potential to drive future research and inform clinical decisions, ultimately contributing to improved strategies for heart disease prevention.

6 Results and Analysis

A partial dependence plot (PDP) illustrates the marginal effect of one or two features on the predicted outcome of a machine learning model. These plots provide insight into whether the relationship between the target variable and a feature is linear-monotonic or more complex.
PDPs are global methods, as they consider all instances to describe the overall relationship between a feature and the predicted outcome. However, these plots rely on the assumption that the distribution of the first feature is independent of the second feature. If this assumption is violated, the calculated averages for the partial dependence plot may include highly unlikely or even impossible data points. Despite this limitation, PDPs remain a valuable tool for interpreting global feature relationships and identifying trends that inform predictive models.

6.1 Feature Analysis and Model Explanation

Features (top): These are the names of the features that contribute to the model's prediction. In this example, the features are "sex" and "age."

Base Value (leftmost bar): This represents the average prediction of the model when none of the features have any influence, serving as a baseline starting point in many cases.

SHAP Values: The colored bars illustrate the contribution of each feature value to the model's output. The color indicates the direction of the impact: red specifies a positive impact (increasing the prediction), while blue specifies a negative impact (decreasing the prediction).

Final Prediction: The final output of the model is obtained by combining the base value with the sum of all SHAP values for the specific data point being analyzed.

6.2 Interpretation of the Plot

- The model appears to predict outcomes related to a population where factors such as "age" and "sex" are influential. For instance, the prediction could pertain to income, health risk, or loan eligibility. - A positive SHAP value for "age" (depicted in red) implies that higher age values increase the predicted outcome. - A negative SHAP value for "sex" (depicted in blue) suggests that belonging

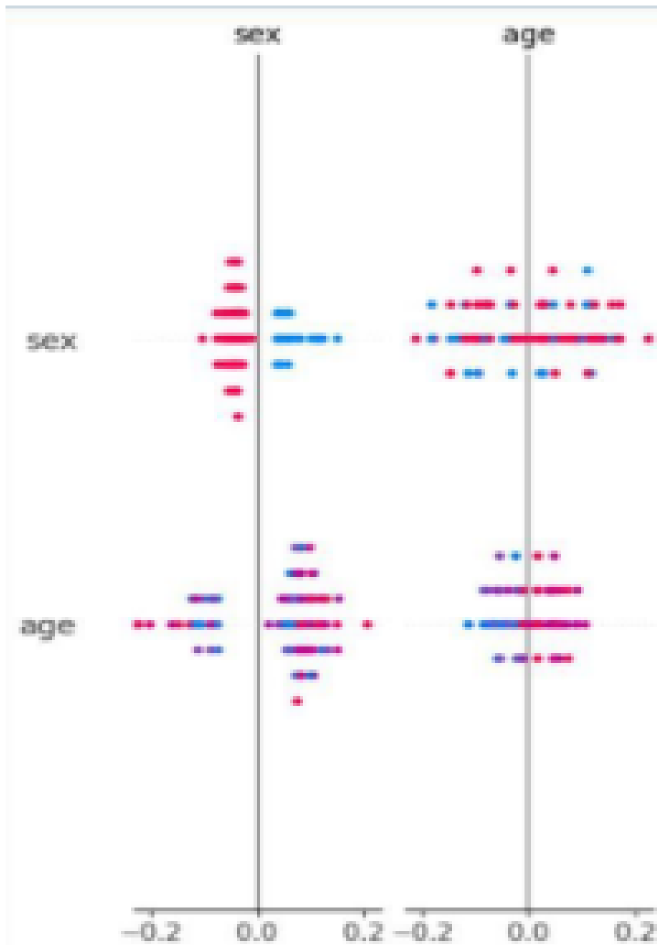

Figure 4: A Correlation Plot Between Sex and Age: Distribution and Interactions

to a specific sex class decreases the predicted value, potentially indicating the presence of a dominant class.

6.3 Dependence Plot

A dependence plot provides insights into how model predictions vary with a single feature. In this case, it demonstrates that the property value significantly increases when the average number of rooms per dwelling exceeds 6.
- Each dot in the plot represents a single prediction (a row in the dataset).
- The x-axis shows the actual feature values from the dataset, while the y-axis represents the SHAP values, indicating the extent to which knowing that

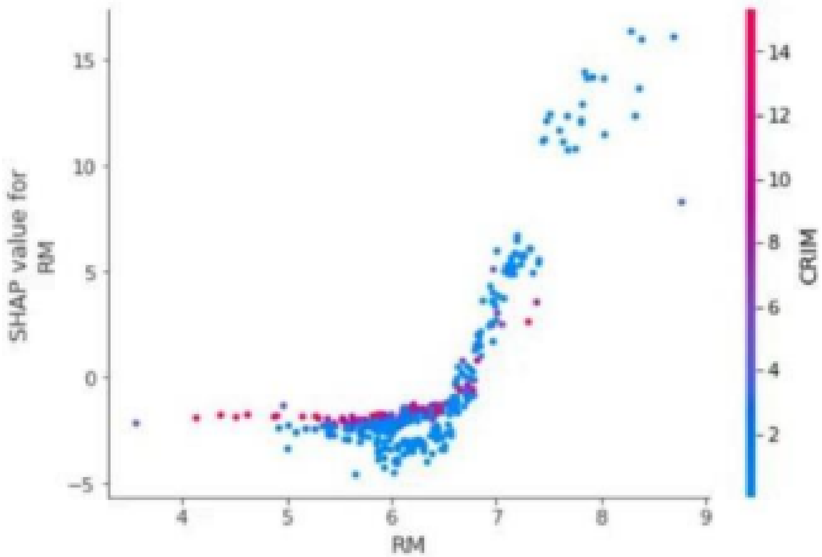

Figure 5: SHAP Beeswarm Plot: Impact of RM on Model Predictions with CRIM Gradient

feature value alters the model's prediction. - The color of the dots corresponds to a secondary feature that may interact with the primary feature. Interaction effects manifest as vertically colored patterns, revealing dependencies between the features.

6.4 SHAP Summary Plot

The `shap.summary_plot()` function generates a visual representation that provides insights into how individual features influence a model's predictions. The plot typically includes the following elements:
- The distribution of SHAP values for each feature, representing both positive and negative contributions to the predictions. - Feature values displayed along the x-axis, offering a clear view of how feature values relate to their impact.
- A dependence plot showing how a specific feature's value interacts with the model's prediction.

This analysis calculates the SHAP values for a machine learning model's predictions on test data. The summary plot is then used to explore the importance of each feature, facilitating a deeper understanding of the model's behavior.

Conclusion And Future Work

The results of the heart attack prediction model demonstrated satisfactory performance, with cross-validation achieving an average accuracy exceeding 85 percentage. Among the evaluated machine learning algorithms—SVM, Random Forest, XGBoost, and k-NN—it was observed that XGBoost and SVM provided the most optimal performance. Decision trees and gradient boosting were also utilized to handle complex interactions within patient data. To ensure interpretability, SHAP and LIME were employed. SHAP analysis highlighted key features such as cholesterol, age, ca, cp, thalach, and oldpeak as critical for risk prediction. Similarly, LIME offered quantitative insights into the differences between model predictions for positive and negative classes, enhancing understanding of both True Positive and True Negative outcomes.

These models enable interpretable and transparent predictions, empowering healthcare providers to make informed decisions. By deconstructing feature contributions, the findings remain patient-specific and closely tied to individual risk factors. The benefits of incorporating SHAP and LIME in this project extend beyond predictive accuracy, offering practical applications for early intervention and effective risk management. These methods establish a foundation for future advancements aimed at improving decision-making in real-world healthcare scenarios and enhancing patient outcomes.

Future work in this field can explore the following directions:

- **Integration with Real-Time Data:** Incorporating real-time patient data from electronic health records (EHRs) or wearable devices could further personalize risk predictions and enable timely interventions.
- **Model Generalization:** Expanding the dataset to include more diverse populations across different demographics, geographies, and medical histories can improve the generalizability of the models.
- **Advanced Interpretability Techniques:** Exploring more advanced interpretability methods, such as counterfactual explanations or causal inference models, could provide deeper insights into the relationships between features and predictions.
- **Clinical Trials and Implementation:** Testing the models in real-world clinical settings and integrating them into decision-support systems to validate their utility and practicality for healthcare providers.
- **Automated Early Warning Systems**: Developing automated alert systems for high-risk patients based on model predictions could assist healthcare teams in proactive management.

- **Multi-Task Learning Models:** Designing models capable of predicting not only heart attack risks but also associated comorbidities or complications, providing a more comprehensive assessment of patient health.

By addressing these areas, the potential of machine learning models to transform healthcare decision-making and improve patient outcomes can be significantly enhanced.

Acknowledgments

We sincerely thank the organizing committee for recognizing our paper and accepting it to be presented at the conference. We would also like to thank the anonymous reviewers who gave useful suggestions, which were taken into account.

References

[1] M. M. Ahsan, S. A. Luna, and Z. Siddique. Machine-learning-based disease diagnosis: A review. *Healthcare*, pages 1–30, 2022.

[2] X. Y. Gao, A. Amin Ali, H. Shaban Hassan, and E. M. Anwar. Improving the accuracy for analyzing heart diseases prediction based on advanced machine learning techniques. 2022.

[3] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, and A. Lamgunde. Heart disease prediction using data mining techniques. In *Proceedings of the 2017 International Conference on Emerging Trends in Engineering and Technology (ICETET)*, 2017.

[4] N. S. Chandra Reddy, S. Shue Nee, L. Zhi Min, and C. Xin Ying. Classification and feature selection approaches by machine learning techniques: Heart disease prediction. *International Journal of Innovative Computing*, 9(1):39–46, 2019.

[5] R. H. Scheuermann, W. Ceusters, and B. Smith. Toward an ontological treatment of disease and diagnosis. In *AMIA Summit on Translational Bioinformatics*, pages 116–120, 2009.