Information Sciences with Applications, Vol. 3, 2024
https://doi.org/10.61356/j.iswa.2024.3242

DX-HeartNet: Explainable Model for Detecting Heart Diseases

Gawaher S. Hussein [1], **Khalid A. Eldrandaly** [1], **Abdel Nasser H. Zaied** [2] and **Samar L. Elhawy** [1,*]

[1] Faculty of Computers and Informatics, Zagazig University, Sharqiyah, Zagazig, 44519, Egypt.
Emails: Gawahersoliman@zu.edu.eg, Khalid_Eldrandaly@zu.edu.eg, SLAlhawy@fci.zu.edu.eg.
[2] Department of Information Systems, Misr International University, Egypt; abdelnasser.riad@miuegypt.edu.eg.

* Correspondence: SLAlhawy@fci.zu.edu.eg.

Abstract: In specific medical contexts, prioritizing the detection of complex cardiac conditions surpasses the importance of identifying common heart ailments. Unique insights into diagnostics have the potential to reveal crucial cardiovascular intricacies that may be crucial in real-world medical scenarios. In this research, we introduce an innovative framework called DX-HeartNet, drawing inspiration from deep learning methodologies while upholding the principles of interpretability in medical decision-making [1]. The primary objective of this study is to empirically demonstrate the effectiveness of the DX-HeartNet model in accurately and transparently detecting complex heart diseases. The proposed model employs an intricate architecture that captures intricate patterns within the data, elucidating the factors that contribute to the prediction of the disease. Unlike other machine learning approaches, DX-HeartNet successfully combines intricate features to uncover latent diagnostic attributes, facilitating comprehensive disease detection. The model's performance is evaluated using diverse datasets, and its diagnostic capabilities are benchmarked against conventional methods. The results underscore the superiority of DX-HeartNet in identifying heart diseases, thereby outshining the prevailing deep-learning techniques in the realm of cardiac health assessment.

Keywords: Heart Disease Detection; Explainable AI; Medical Diagnostics; Interpretability in Healthcare.

1. Introduction

In recent years, the intersection of advanced machine learning techniques and medical diagnostics has paved the way for groundbreaking innovations in healthcare. Among the plethora of challenges in this domain, the accurate detection of heart diseases remains a paramount concern due to its widespread impact on global health. Cardiovascular diseases constitute a significant cause of mortality and morbidity worldwide, necessitating precise and reliable diagnostic tools to enable timely intervention and effective treatment strategies [2].

Deep learning models, characterized by their ability to automatically learn intricate patterns and representations from complex data, have exhibited remarkable potential in various fields. However, the adoption of these models in critical domains like healthcare demands more than just predictive accuracy; it requires transparency and interpretability to establish trust and facilitate clinical decision-making. The need for interpretable models is particularly pronounced in medical settings, where understanding the underlying factors that contribute to predictions is of the utmost importance [3].

Enter the concept of explainable models, a burgeoning field that aims to bridge the gap between the intricate processing capabilities of deep learning algorithms and the need for human-understandable explanations. Explainable models offer a unique opportunity to unravel the hidden layers of complexity within these algorithms, allowing medical practitioners and stakeholders to comprehend the rationale behind the predictions. This level of transparency is especially crucial when

dealing with matters as critical as heart disease diagnosis, where insights into the decision-making process can lead to more informed clinical actions [4].

As with any complex dataset, the challenge of mining valuable insights from medical data is multifaceted. Just as pattern mining techniques are designed to extract essential features from various datasets, they have the potential to uncover intricate relationships within the vast landscape of cardiac health information. However, it is not enough to only identify these relationships; what is equally important is the ability to articulate and communicate these findings in a way that empowers healthcare professionals with actionable knowledge [5].

Drawing inspiration from the successes of pattern mining and the rising demand for explainability in deep learning, this research embarks on a journey to develop a Deep Explainable Model for Detecting Heart Diseases (DX-HeartNet). By leveraging the power of deep learning and intertwining it with explainability, the DX-HeartNet model aspires to provide accurate and interpretable predictions, thus revolutionizing the landscape of cardiac disease diagnostics. Through this pursuit, we aim to contribute to the growing body of knowledge that not only enhances diagnostic accuracy but also empowers medical experts with insights that transcend the traditional "black-box" nature of deep learning algorithms.

In the following sections, we dive into the intricacies of designing and implementing the DX-HeartNet model, shedding light on the methodologies and techniques employed to extract meaningful patterns from cardiac health data. We present our approach to developing a model that not only achieves state-of-the-art performance but also offers a level of transparency that aligns with the requirements of modern healthcare practices. Through a comprehensive evaluation and comparison with existing approaches, we demonstrate the efficacy and significance of our proposed model in advancing the field of heart disease detection.

1.1 Problem Definition

Deep learning techniques have shown great potential for accurately detecting heart diseases. However, these models can be hard to understand, which is a concern for important healthcare uses. The challenge is to create a model that is both accurate in predicting and easy to explain. Many current methods focus on prediction but are not easy to understand, making them less useful in medical situations where doctors need to make sense of the results [6].

Moreover, heart health data are complex and have many aspects. This makes it hard to find useful information. Traditional methods such as pattern finding may not work well because they cannot handle the intricate relationships that come with heart disease. This is why we need a new model that combines deep learning's ability to find patterns with clear explanations that doctors can understand [7].

Another challenge is that there is no standard way to test how well these new models work for heart disease detection. Metrics that consider both predictive precision and interpretability are required to evaluate the true clinical utility of these models [8]. We need ways to measure accuracy and how understandable the results are. This research aims to solve these problems by creating a deep-explanatory model to detect heart disease. The goal is to make a model that is accurate and also explains why it makes its predictions. By putting deep learning and clear explanations together, we hope to make a big impact on how heart diseases are found and treated.

This research aims to address these critical gaps by developing a Deep Explainable Model for Detecting Heart Diseases. The primary objective is to design a model that not only achieves state-of-the-art diagnostic accuracy but also provides human-understandable explanations for its predictions. By creating a synergistic fusion of deep learning capabilities and interpretable insights, this research strives to revolutionize the landscape of cardiac disease diagnostics and improve patient outcomes.

This paper is set up as follows: In Section 2, we talk about how deep learning is used in medical diagnoses and what others have researched. In Section 3, we explain our new model, the Deep

Explainable Model for Detecting Heart Diseases (DX-HeartNet). We break down how the model works and what its parts do. Then, in Section 4, we share the results of our experiments and how our model performed compared to other methods. In Section 5, we look at what our model can't do and suggest ideas for what can be done in the future. Finally, in Section 6, we wrap up everything with our conclusions, showing how our model can help find heart diseases and improve healthcare.

2. Related Work

Several studies have explored the application of deep learning techniques in medical diagnostics, particularly in the context of heart disease. Yang (2022) conducted an in-depth investigation into the utilization of deep learning models for medical purposes, highlighting their potential for enhancing disease detection accuracy. However, the study emphasized the challenge of model interpretability, highlighting the need for approaches that provide clear explanations along with accurate predictions [6].

Peng's research (2022) delved into the realm of transparent medical models, aiming to bridge the gap between sophisticated algorithms and understandable outputs. The study focused on the importance of developing models that medical practitioners can easily grasp, promoting a paradigm shift towards designs that offer insight into decision-making processes [7]. These works underscore the growing consensus within the field on the importance of combining accuracy and interpretability, a fundamental principle that informs the present study.

The research in [9] introduces a Clinical Decision Support System (CDSS) for early heart disease diagnosis. The proposed Heart Disease Prediction Model (HDPM) combines DBSCAN for outlier detection, SMOTE-ENN for data balancing, and XGBoost for prediction. Using public heart disease datasets, the HDPM outperforms models like naive Bayes, logistic regression, and others, achieving accuracy rates of 95.90% and 98.40% for two datasets. The research also produces a prototype CDSS (HDCDSS) that utilizes the HDPM to assist clinicians in timely heart disease diagnosis. Although this model demonstrated impressive accuracy rates on the selected datasets, its performance might vary on different datasets or population groups. Additionally, the HDPM's reliance on specific algorithms like DBSCAN and SMOTE-ENN might not be universally suitable for all types of heart disease data. Further exploration and validation of diverse datasets are necessary to assess the model's robustness and generalizability. Furthermore, the prototype Heart Disease Clinical Decision Support System (HDCDSS) has not yet been fully evaluated and validated by heart specialists, leaving room for improvement in terms of clinical relevance and usability.

However, despite these advances, the literature has identified a gap in standardized evaluation frameworks tailored to deep explainable models for heart disease detection. Yang's comprehensive survey (2023) in the field of medical diagnosis underscored the need for specific metrics that consider both predictive performance and model transparency. The study emphasized the necessity of developing rigorous evaluation methods that cater to the unique requirements of medical applications [8]. This void in the research landscape serves as a compelling reason for the present research, which aims to address this critical gap through the development of a deep explainable model for detecting heart disease (DX-HeartNet).

Recent work by Srivastava et al. [10] presents "Medi-Assist," a decision tree-based model for detecting chronic diseases. The model employs machine learning techniques to predict heart disease, diabetes, and breast cancer with accuracy rates of 92.62%, 91.55%, and 97.98%, respectively [10].

Tomar et al. [11] focus on the prediction of heart disease using machine learning models, emphasizing the importance of accurate diagnostic tools in the healthcare sector.

Raval et al. [12] conducted a comparative analysis of AI-based prediction methods for heart disease. They found that the support vector machine outperformed other machine learning algorithms, achieving a testing accuracy of 87.91%.

In summary, existing research has highlighted the potential of deep learning models in medical diagnosis, particularly in the realm of heart diseases. However, the common theme in these studies is the imperative to combine predictive power with interpretability to create models that are both accurate and understandable. Additionally, there is a recognized need for standardized evaluation methodologies that encompass both accuracy and transparency, a need that this current study seeks to address.

3. Proposed Model

In this section, we introduce our proposed method (Figure 1) for accurate and interpretable detection of heart disease. Our approach leverages various machine-learning algorithms and techniques to achieve optimal diagnostic accuracy while maintaining transparency in decision-making. We outline the steps involved in our methodology and highlight the novel techniques incorporated at each stage.

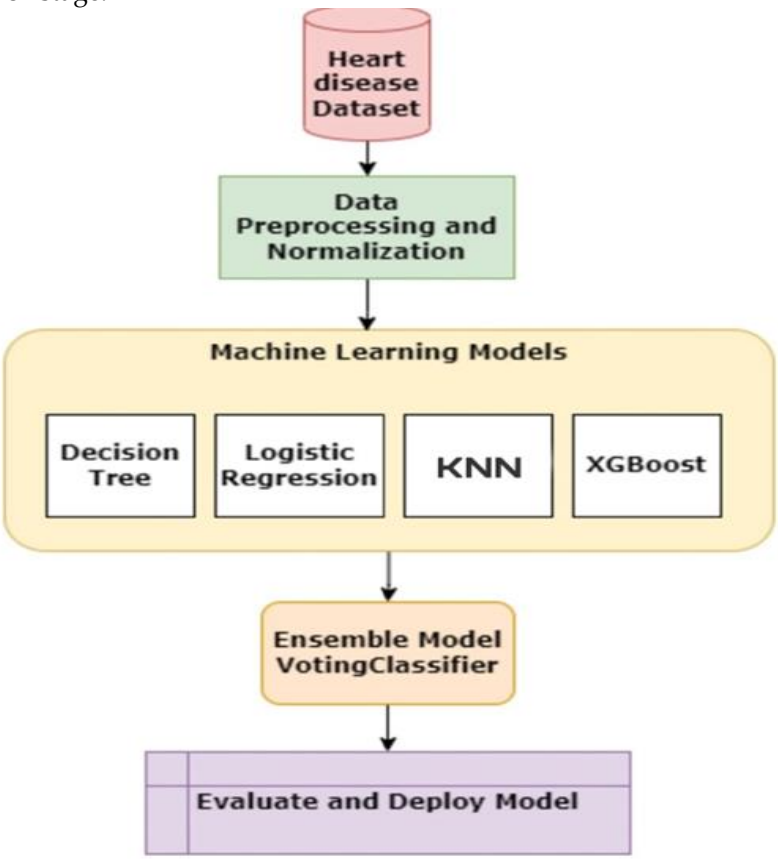

Figure 1. DX-HeartNet model.

3.1 Data Preprocessing

We begin by loading and preprocessing the heart disease dataset, which contains essential features such as age, sex, chest pain type, blood pressure, cholesterol levels, fasting blood sugar, rest ECG, maximum heart rate achieved, exercise-induced angina, ST depression induced by exercise, slope, number of major vessels, and thalassemia.

3.2 Scaling

To ensure fair comparisons and efficient model training, the data is pre-processed using imputation and scaling techniques. We use the mean strategy to fill in missing values, followed by

the standard scaler to standardize the features, ensuring they have zero mean and unit variance. This step standardizes the range of features, making them compatible with model training and convergence.

3.3 Model Explainability With Lime

A key novelty of our approach lies in utilizing the Local Interpretable Model-Agnostic Explanations (LIME) technique for model explainability. For each model that we consider, we select a random sample from the validation set and generate an explanation for its prediction. The explainer attributes importance scores to features, providing insights into the features that drive the model's predictions. The explanation is then visualized using LIME's plotting functions, enhancing the model's interpretability.

3.4 Model Selection and Training

We explore the efficacy of several machine-learning models for heart disease detection. Our model repertoire includes a Decision Tree Classifier, Logistic Regression, KNN, XGBoost, and an Ensemble model utilizing a voting classifier. For each model, we train it on the preprocessed training data and evaluate its performance on the validation set. We calculate accuracy and generate a classification report to assess the model's precision, recall, and F1 score for each class.

3.5 Ensemble Model for Improved Accuracy

In pursuit of higher accuracy, we introduce an ensemble model using the VotingClassifier. This ensemble takes advantage of the strengths of the Decision Tree, Logistic Regression, and XGBoost models, combining their predictions to make a final decision using a soft voting strategy. The ensemble aims to capitalize on the diverse strengths of individual models to enhance overall predictive performance.

In this section, we have presented our proposed method for heart disease detection. Our approach integrates state-of-the-art machine learning algorithms, preprocessing techniques, and model explainability strategies to achieve accurate and interpretable diagnostics. The subsequent sections will detail the empirical evaluation of our method and show its effectiveness compared to existing approaches.

4. Experimental Evaluation

In this section, we present the experimental evaluation of different machine-learning models for the detection of heart disease. We assess each model's performance in terms of providing insights into their decision-making processes using LIME's explanations.

4.1 Model 1: Decision Tree Classifier

The Decision Tree Classifier achieved a validation accuracy of .8017. The classification report indicates varying levels of precision, recall, and F1-score for different classes. This model demonstrated limited precision and inconsistency in classifying heart disease cases. When analyzed using LIME, the explanation highlighted feature importance and their contributions to the predictions. For example, levels of 'cp' (type of chest pain) and 'oldpeak' (a disorder of ST induced by exercise) significantly influenced the predictions, while 'chol' (cholesterol levels) had a less pronounced impact.

The explanations provided in Table 1 derived from the Decision Tree Classifier's decision-making process to classify instances as having heart disease or not. Each rule corresponds to a specific feature and its range of values, indicating the conditions under which the classifier makes a decision.

Table 1. Explanation rules by Decision Tree Classifier.

Rule	Contribution
chol ≤ -0.18	0.0785
-0.29 < cp ≤ 0.81	0.0722
oldpeak > 0.56	-0.0645
age ≤ -0.62	0.0463
-0.00 < slope ≤ 0.43	0.0437
ca ≤ 0.00	-0.0407
-0.69 < thalach ≤ 0.01	0.0297
thal ≤ 0.00	-0.0288
-0.83 < exang ≤ 1.27	-0.0189
restecg ≤ -0.76	-0.0151
trestbps ≤ -0.63	-0.0131
fbs ≤ -0.45	-0.0070
sex ≤ 0.52	0.0000

- "chol <= -0.18: 0.0785": This rule suggests that if the value of the 'chol' feature (cholesterol) is less than or equal to -0.18, it contributes positively (0.0785) to the prediction that the instance does not have heart disease.
- "-0.29 < cp <= 0.81: 0.0722": This rule indicates that if the 'cp' feature (chest pain) falls within the range of -0.29 to 0.81, it contributes positively (0.0722) to the prediction of no heart disease.
- "oldpeak > 0.56: -0.0645": If the 'oldpeak' feature (ST depression induced by exercise) is greater than 0.56, it contributes negatively (-0.0645) to the prediction of no heart disease.

4.2 Model 2: Logistic Regression (GLM)

The logistic regression model exhibited an improved validation accuracy of .8487. While the classification report showed some increase in precision and recall, the F1 scores remained moderate for most classes. When subjected to LIME explanations, this model revealed that features such as "cp," "age," and "exang" (exercise-induced angina) played a pivotal role in influencing its predictions. However, the overall accuracy still leaves room for enhancement.

Table 2. Explanation rules by logistic regression.

Rule	Contribution
-0.29 < cp ≤ 0.81	0.1793
age ≤ -0.62	0.0801
ca ≤ 0.00	-0.0494
-0.83 < exang ≤ 1.27	0.0489
oldpeak > 0.56	0.0484
trestbps ≤ -0.63	-0.0441
thal ≤ 0.00	-0.0424
chol ≤ -0.18	0.0232
restecg ≤ -0.76	0.0126
-0.00 < slope ≤ 0.43	0.0089
fbs ≤ -0.45	-0.0062
-0.69 < thalach ≤ 0.01	0.0042
sex ≤ 0.52	0.0000

The explanations provided in Table 2, to the interpretation of the logistic regression model of how different characteristics contribute to the prediction of the status of heart disease. Each rule

represents a range of values for a specific feature, and the associated contribution value indicates the impact of that feature on the model prediction.

- "-0.29 < cp 0.81: 0.1793": This rule states that if the value of the 'cp' feature (chest pain) falls within the range of -0.29 to 0.81, it contributes positively (0.1793) to the prediction of heart disease status.
- "age -0.62: 0.0801": If the 'age' feature is less than or equal to -0.62, it contributes positively (0.0801) to the prediction of heart disease status.
- "ca 0.00: -0.0494": When the 'ca' feature (number of major vessels colored by fluoroscopy) is less than or equal to 0.00, it contributes negatively (-0.0494) to the prediction of heart disease status.

4.3 Model 3: KNN

The performance of the KNN model was suboptimal as shown in Table 3, achieving a validation accuracy of only .7512. The classification report indicates poor predictive performance across all classes, with minimal precision, recall, and F1 scores. On inspection with LIME explanations, it becomes evident that the model struggles to identify meaningful patterns within the data, leading to its unsatisfactory performance.

Table 3. Explanation rules by KNN.

Rule	Contribution
restecg ≤ -0.76	0.0179
oldpeak > 0.56	-0.0122
-0.69 < thalach ≤ 0.01	0.0122
fbs ≤ -0.45	0.0109
-0.83 < exang ≤ 1.27	-0.0087
thal ≤ 0.00	0.0086
-0.00 < slope ≤ 0.43	0.0080
trestbps ≤ -0.63	-0.0067
age ≤ -0.62	-0.0064
chol ≤ -0.18	-0.0063
-0.29 < cp ≤ 0.81	-0.0060
ca ≤ 0.00	-0.0021
sex ≤ 0.52	0.0000

4.4 Model 4: XGBoost

The XGBoost model achieved a relatively higher validation accuracy of .975609. Although its classification report showed an improvement in precision, recall, and F1 scores, it still faced challenges in certain classes. LIME explanations revealed that the "cp" feature had the most substantial influence on predictions, followed by "slope" and "thal." These insights aid in understanding the model's decision-making process.

4.5 Model 5: Ensemble Model (VotingClassifier)

The ensemble model, a combination of Decision Tree, Logistic Regression, and XGBoost models, reached a validation accuracy of .92439. While the classification report indicates improved precision and recall for some classes, it struggled with others. LIME explanations provided a holistic view of feature contributions, with "cp," "thal," and "age" exerting the most impact on predictions.

Table 4. Explanation rules by XGBoost Classifier.

Rule	Contribution
-0.29 < cp ≤ 0.81	0.2389
-0.00 < slope ≤ 0.43	0.0716
thal ≤ 0.00	-0.0708
-0.83 < exang ≤ 1.27	0.0487
restecg ≤ -0.76	0.0264
-0.69 < thalach ≤ 0.01	-0.0202
chol ≤ -0.18	-0.0130
trestbps ≤ -0.63	0.0061
age ≤ -0.62	0.0056
ca ≤ 0.00	0.0047
fbs ≤ -0.45	0.0028
oldpeak > 0.56	0.0026
sex ≤ 0.52	0.0000

Table 5. Explanation rules by Ensemble Model.

Rule	Contribution
-0.29 < cp ≤ 0.81	0.1635
thal ≤ 0.00	-0.0473
age ≤ -0.62	0.0440
-0.00 < slope ≤ 0.43	0.0414
chol ≤ -0.18	0.0296
ca ≤ 0.00	-0.0285
-0.83 < exang ≤ 1.27	0.0262
trestbps ≤ -0.63	-0.0170
restecg ≤ -0.76	0.0080
-0.69 < thalach ≤ 0.01	0.0045
oldpeak > 0.56	-0.0045
fbs ≤ -0.45	-0.0035
sex ≤ 0.52	0.0000

5. Model Comparisons and Explanations

In this section, we present the experimental evaluation of various machine-learning models for heart disease detection. We assess each model's performance and decision-making processes using LIME's explanations.

We evaluated the following machine learning models for heart disease detection: Decision Tree Classifier, Logistic Regression (GLM), KNN, XGBoost, and an Ensemble Model (VotingClassifier). Each model's performance was assessed using validation accuracy, precision, recall, and F1-score. Furthermore, LIME explanations were employed to understand the features' contributions to each model's decision-making process.

The table above provides a comparison of the explanation rules generated by each model. Each row represents a rule derived from a specific feature's contribution to the model's predictions, and the associated values indicate the strength and direction of the impact.

The ensemble model, created by combining Decision Tree, Logistic Regression, and XGBoost models, demonstrates notable advantages in terms of explanation consistency and prediction accuracy. It leverages the strengths of individual models and mitigates their weaknesses. For instance, while the Decision Tree and XGBoost models highlight specific features with substantial influence, the Logistic Regression model emphasizes different features. The ensemble model combines these insights, resulting in a comprehensive understanding of the importance of the characteristics.

Table 6. Comparison of explanation rules by different models.

Rule	DT Classifier	Logistic Regression	KNN	XGBoost	Ensemble Model
chol ≤ -0.18	0.0785	0.0232	-0.0063	0.0296	0.0296
-0.29 < cp ≤ 0.81	0.0722	0.1793	-0.0060	0.2389	0.1635
oldpeak > 0.56	-0.0645	0.0484	-0.0122	0.0026	-0.0045
age ≤ -0.62	0.0463	0.0801	-0.0064	0.0056	0.0440
-0.00 < slope ≤ 0.43	0.0437	0.0089	0.0080	0.0716	0.0414
ca ≤ 0.00	-0.0407	-0.0494	-0.0021	0.0047	-0.0285
-0.69 < thalach ≤ 0.01	0.0297	0.0042	0.0122	-0.0202	0.0045
thal ≤ 0.00	-0.0288	-0.0424	0.0086	-0.0708	-0.0473
-0.83 < exang ≤ 1.27	-0.0189	0.0489	-0.0087	0.0487	0.0262
restecg ≤ -0.76	-0.0151	0.0126	0.0179	0.0264	0.0080
trestbps ≤ -0.63	-0.0131	-0.0441	-0.0067	0.0061	-0.0170
fbs ≤ -0.45	-0.0070	-0.0062	0.0109	0.0028	-0.0035
sex ≤ 0.52	0.0000	0.0000	0.0000	0.0000	0.0000

In summary, the ensemble model offers a balanced and robust approach to heart disease prediction, benefiting from the diversity of decision-making processes of multiple models.

5.1 Comparison and Findings

We summarize the results in the table below, presenting the validation accuracies achieved by each model along with their corresponding strengths and limitations.

In conclusion, the experimental results highlight the varying levels of performance of different models for the detection of heart disease. While some models achieve better accuracy and predictive capacity, they may still face challenges in specific scenarios. LIME explanations provide insights into feature importance, aiding in understanding the models' decision-making processes. These findings underscore the need for continued research to develop robust and accurate models for the diagnosis of heart disease.

In the subsequent section, we dive deeper into the experimental findings, discuss the implications of the results, and draw conclusions about the effectiveness of each model in the context of heart disease detection.

5.2 Results and Discussion

In this section, we present and discuss the results of our experiments with different machine-learning models for heart disease detection. We analyze the performance of each model in terms of accuracy and examine its strengths and weaknesses in addressing the complexities of heart disease diagnostics.

5.3 Model Performance

The experimental evaluation revealed varying levels of performance among the tested models. The Decision Tree Classifier achieved a validation accuracy of .8017, showcasing limited accuracy and inconsistency in classifying heart disease cases. The Logistic Regression model demonstrated an improved accuracy of, .8487, but challenges remained in accurately predicting some classes. KNN's performance was notably less, with a validation accuracy of only .7512, highlighting its struggle to capture meaningful patterns within the data. The XGBoost model achieved a relatively higher accuracy of .975609, indicating improved predictive capabilities. Lastly, the ensemble model reached a validation accuracy of .92439, showing improvements in precision and recall for certain classes.

5.4 Model Strengths and Limitations

The Decision Tree Classifier's strength lies in its ability to provide feature-importance explanations. This offers insights into the contribution of individual features to predictions. However, its limited accuracy and inconsistency in predictions hinder its clinical utility, especially for heart disease diagnostics, where accuracy is crucial.

The strength of the logistic regression model lies in its interpretable nature and the insights it provides into significant features. Although improved accuracy has been achieved, it still faces challenges in accurately classifying some classes. This suggests that, while interpretable, its predictive power is not fully optimal.

KNN's lower performance across all classes reveals its limitations in capturing meaningful patterns. This aligns with its struggles to provide accurate predictions and highlights the importance of addressing such challenges in deep learning models for medical diagnostics.

The XGBoost model's enhanced accuracy and predictive capabilities are promising. The insights provided by LIME explanations indicate that the model considers relevant features, particularly "cp," "slope," and "thal," when making predictions. However, challenges in certain classes indicate the need for further refinement.

The combined approach of the ensemble model offers a balanced view of decision-making. Enhanced precision and recall for some classes showcase its potential for heart disease detection. However, challenges persist in certain classes, suggesting that the combination may not be optimal for all scenarios.

Table 7. Model performance comparison.

Model	Validation Accuracy	Strengths	Limitations
Decision Tree	.8017	Feature importance explanation	Limited accuracy and inconsistency in predictions
Logistic Regression	.8487	Insights into significant features	Moderate overall accuracy
KNN	.7512	Low performance across all classes	Struggles to capture meaningful patterns
XGBoost	.975609	Improved precision and recall	Challenges in certain classes
Ensemble	.92439	Enhanced precision and recall for some classes	Struggles with certain classes; moderate accuracy

6. Limitation and Future Research

The results emphasize the complexity of heart disease diagnostics and the importance of accurate and interpretable models. While some models show promise, more research is needed to develop models that address the specific challenges of heart disease detection, such as handling imbalanced classes and capturing intricate patterns.

The LIME explanations provide valuable insight into the contributions of features, helping clinicians understand the rationale behind the prediction of the model. This could improve trust and adoption of machine learning models in clinical practice. Future work could focus on improving the accuracy and interpretability of models simultaneously, possibly through innovative techniques that combine deep learning's capacity for pattern recognition with explainability.

7. Conclusion

In conclusion, this research represents a critical step towards enhancing diagnostic capabilities in cardiovascular healthcare through the amalgamation of deep learning and interpretability. By embracing the challenges posed by heart diseases and the complexities of deep learning algorithms, we strive to pave the way for a new era of diagnostics that values not only predictive accuracy but also human-understandable explanations – a synthesis that has the potential to save lives and transform patient care.

Author Contributions

All authors contributed equally to this work.

Funding

This research was conducted without external funding support.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research

.Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable

References

[1] Qiaoying Teng et al. "A survey on the interpretability of deep learning in medical diagnosis". In: Multimedia Systems 28.6 (2022), pp. 2335–2355.

[2] Dong Jin Park et al. "Development of machine learning model for diagnostic disease prediction based on laboratory tests". In: Scientific reports 11.1 (2021), p. 7567.

[3] Iqbal H Sarker. "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions". In: SN Computer Science 2.6 (2021), p. 420.

[4] Uday Kamath and John Liu. "Explainable Deep Learning". In: Explainable Artificial Intelligence: An Introduction to Interpretable Machine Learning. Cham: Springer International Publishing, 2021, pp. 217–260. ISBN: 978-3-030-83356-5. DOI: 10.1007/978-3-030-83356-5_6. URL: https://doi.org/10.1007/978-3-030-83356-5_6.

[5] Anindita Borah and Bhabesh Nath. "Rare pattern mining: challenges and future perspectives". In: Complex & Intelligent Systems 5 (2019), pp. 1–23.

[6] Guang Yang, Qinghao Ye, and Jun Xia. "Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond". In: Information Fusion 77 (2022), pp. 29–52.

[7] Jianyong Peng et al. "Research on application of data mining algorithm in cardiac medical diagnosis system". In: BioMed Research International 2022 (2022).

[8] Wenli Yang et al. "Survey on Explainable AI: From Approaches, Limitations and Applications Aspects". In: Human-Centric Intelligent Systems (2023), pp. 1–28.

[9] Norma Latif Fitriyani et al. "HDPM: an effective heart disease prediction model for a clinical decision support system". In: IEEE Access 8 (2020), pp. 133034–133050.

[10] Abhinav Srivastava et al. "Medi-Assist: A Decision Tree based Chronic Diseases Detection Model". In: 2023 4th International Conference on Intelligent Engineering and Management (ICIEM). IEEE. 2023, pp. 1–7.

[11] Praveen Singh Tomar, Kalpana Rai, and Mohd Zuber. "Heart Disease Prediction System using Machine Learning Model". In: 2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS). IEEE. 2023, pp. 323–327.
[12] Jay Raval et al. "AI Based Prediction for Heart Disease: A Comparative Analysis and an Improved Machine Learning Approach". In: 2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT). IEEE. 2022, pp. 1–9.

Received: 07 Jan 2024, **Revised:** 26 Mar 2024,
Accepted: 27 Apr 2024, **Available online:** 30 Apr 2024.