www.ijecs.in
International Journal Of Engineering And Computer Science
Volume 13 Issue 09 September 2024, Page No. 26391-26400
ISSN: 2319-7242 DOI: 10.18535/ijecs/v13i08.4887

Stacking Ensemble Artificial Intelligence Model for Heart Disease Diagnosis

Njoku Felix Anayo [1]. , Frank Edughom Ekpar [2].

[1,] Topfaith University, Nigeria..

[2] Scholars University Ltd; Rivers State University; Topfaith University, Nigeria.

Abstract
Personalized treatment plans, predictive analytics, and artificial intelligence (AI)-driven diagnostics are becoming more and more popular as a way to improve decision-making, expedite operations, and improve patient care. But there are still a number of substantial barriers to overcome, which includes but not limited to issues with user adoption, trust, prejudice, and fairness brought on by resistance from healthcare providers and a lack of confidence in the system's recommendations. To overcome these challenges and realize the full potential of AI-driven solutions, the system's accuracy and safety through meticulous testing and validation of AI algorithms becomes indispensable. This research provides a hybrid AI model that blends three base models with a Meta model in order to diagnose heart disease effectively. The essence is to revalidate the existing AI diagnostic models for cardiac diseases diagnostics and to tackle the concerns impeding the full utilization of the available AI diagnostic system. The study makes use of each model's advantages by merging these various and complimentary algorithms into a stacking ensemble model to create a diagnostic system that is more potent. Using publicly available heart disease data, the model performs remarkably well; it achieves 89% accuracy, 85% recall (sensitivity), 92% specificity, and 89% precision. This hybrid model's performance and proven efficacy are expected to boost trust in the system's recommendations and encourage broader implementation in clinical practice.

Key Words: Stacking Ensemble, Artificial Intelligence (AI), Heart Disease, Diagnosis, AI-driven Diagnostics, Hybrid AI Model, healthcare.

1. Introduction

Artificial Intelligence (AI) is hastily developing and becoming a key component of the healthcare landscape by providing cutting-edge diagnostic tools, individualized treatment programs, and predictive analytics. These AI-powered technologies have the golden touch to improve decision-making processes, expedite operations, and improve patient care [1], [2]. Though AI holds enormous potential for the healthcare industry, a variety of barriers prevent its broad use. These hitches include concerns with user acceptance, trust, bias, and fairness that are mostly brought about by rebuff from healthcare providers and a lack of trust in the system's recommendation [3], [4], [5], [6]. On the other hand, healthcare providers' disapproval is one of the main impediments to the adoption of healthcare IT solutions. This averseness may be due to a number of reasons, such as a fear of change, a lack of technical expertise, or worries about disruptions to plan of action. It can also be perceived that healthcare professionals may be resistant to these new technologies because they see them as a threat to their professional autonomy and also having the potential to increase their burden according research [7], [8]. Furthermore, many practitioners find the switch from conventional paper-based systems to electronic health records (EHRs) intimidating, and that makes them reluctant to embrace new technologies. However, for software solutions to be used and implemented effectively, it is essential for it to garner the requisite trust and inclusive acceptance. The concerns raised by healthcare practitioners bothers on the system's accuracy and dependability, particularly when it comes to hermetic decision-making process. These healthcare providers are more likely to believe recommendations made by human specialists than those generated by algorithms [3], [9]. This mistrust of algorithmic decision-making may prevent healthcare software from being widely used.

Algorithm bias in healthcare is another stern concern. The data used to train these algorithms may have biases because it may not be entirely evocative of the population. According to [10], racial prejudice in an algorithm exploited widely in the US healthcare system resulted in less treatment being given to black patients relative to white patients with equivalent health profiles. These biases raise questions about fairness and ethics in addition to destabilizing public confidence in the system [11], [12]. Subsequently, biased algorithms have the potential to worsen already-existing health disparities, it is imperative that healthcare software solutions should ensure fairness. According to [13] , machine learning models which are not properly created and administered, may even underpin and amplify societal biases [14]. In order to avoid prejudice against any patient group, algorithms must be continuously evaluated and adjusted for fairness especially in healthcare sector due to its connection to life. Research has shown that healthcare providers believe the algorithms don't grasp context, hence, they muster distrust for the system's recommendations [15]. A wide range of contextual considerations that may not be fully captured by algorithms are taken into account when making decisions by humans in the healthcare industry. This obviously raise doubts about the system's capacity to provide accurate and distinct advice. Current advancement in a research, especially explainability in AI systems has delivered ground breaking evidence crucial enough to foster confidence and trust in the healthcare profession [16], [17], [18], [19].

The need for software-based solutions to improve accuracy, timeliness, and reliability in the diagnosis of complicated diseases such as heart disease is apparent. These technologies can greatly enhance patient care and diagnostic results by utilizing cutting-edge algorithms and real-time data processing. The capability of the systems to analyze vast amounts of data more precisely than human practitioners makes it more viable to consider. A deep neural network, for example, can detect arrhythmias from ECG signals with high accuracy, frequently exceeding skilled cardiologists, according to a studies by [20], [21], [22], [23]. These networks can detect minute patterns and correlations that the human eye might overlook by studying large datasets, which enables them to make more precise diagnoses. More so, analysis times can be meaningfully lowered by adopting artificial intelligence (AI) to interpret imaging data. From studies, Artificial intelligence (AI) in echocardiography permitted for faster and more accurate evaluations of heart function than were possible with orthodox techniques [24], [25], [26]. This promptness at which analysis is performed can expedite the diagnosis procedure, enabling earlier intervention and therapy.

Artificial intelligent models can provide reliable performance without the subjectivity or weariness that could tilt the finding of human diagnosticians. For instance, an AI system showed great reliability and consistency in diagnosing diabetic retinopathy, a condition carefully associated to cardiovascular health on the use of deep learning in retinal disease screening [27], [28], [29]. This consistency guarantees that diagnoses are not only prompt and precise but also repeatable and steady over time. In real-world applications and case studies, software-based solutions in cardiac disease diagnostics has demonstrated a number of key advantages. An AI model created by Stanford University researchers could identify pneumonia from chest X-rays more accurately than radiologists. This model indirectly aids cardiovascular diagnostics by identifying associated issues early, given that pneumonia can exacerbate heart diseases [30], [31], [32], [33]. In a similar demonstration, a machine learning-enhanced electrocardiogram (ECG) system at the Mayo Clinic showed great ability to forecast atrial fibrillation (AF) in individuals with normal sinus rhythm. In line with these advances, it is manifest that AI system's predictive capability is apparent for averting strokes and other heart-related issues, proving its dependability and usefulness in clinical settings [34], [35]. Hence, artificial intelligence-enabled wearables, like smartwatches, can continually monitor heart rate and rhythm and instantly notify consumers and medical professionals of potential problems like atrial fibrillation. The prompt and accurate benefits of these software solutions in regulating heart health are exemplified by this continuous monitoring and fast alerting system [36], [37], [38].

Discussed so far is the need to ensure trust, acceptance, and fairness in the adoption and use of use of AI-based diagnostic systems. This paper presents a hybrid AI model that makes use of a stacking ensemble technique in order to address the requirement for accurate AI models in the detection of heart disease. By combining three foundational models, this approach makes use of their complimentary and varied characteristics to produce a more reliable diagnostic system. Impressive performance metrics have been demonstrated by the stacking ensemble model suggested in this paper, which has been trained, tested, and verified on publically available heart disease data. By highlighting the value of thorough validation and the potential advantages of ensemble modeling techniques in obtaining high diagnostic accuracy and reliability, we hope to contribute to the ongoing efforts to improve AI-based diagnostic systems.

2. Materials and Methods

The development and evaluation of a stacking ensemble artificial intelligence model for the detection of cardiac disease is presented in this study. The model leverages the strengths of multiple machine learning algorithms to improve prediction accuracy and dependability. The approach began with the careful gathering and preprocessing of data from multiple sources, including the Cleveland Heart Disease dataset from the UCI Machine Learning Repository. To ensure data integrity and improve model resilience, the dataset was carefully standardized using the min-max normalization technique. Following preprocessing, the dataset comprised of 297 rows, each containing 14 columns. The first 13 columns represented clinical measurements for each participant, including resting blood pressure, serum cholesterol, fasting blood sugar, and resting electrocardiographic (ECG) results, among others. The fourteenth (target) column indicated the diagnosis, where a value of 0 signified normal heart function, and values 1, 2, 3, and 4 indicated varying degrees of heart disease. The diagnosis was modeled as a binary classification problem, with the target column rescaled to a fixed range [0, 1]. This simplifies the problem to a binary classification task.

3. Model Development and Implementation

The core of this study involved constructing a hybrid AI model using a stacking ensemble method. The model integrated three diverse base algorithms such as XGBoost, Random Forest, and Support Vector Machine (SVM) with each selected for its proven efficacy in handling structured medical data. The Random Forest algorithm provides robustness by aggregating the predictions of multiple decision trees, reducing overfitting and enhancing generalization tasks [39], [40] , the Support Vector Machine (SVM) delivers its effectiveness in high-dimensional spaces and precision in handling classification, and the XGBoost algorithm affords its high performance in handling structured data and its ability to capture complex relationships through gradient boosting [41], [42]. The stacking ensemble approach capitalized on the complementary strengths of these models, with a logistic regression meta-learner combining their predictions to maximize overall predictive performance as depicted by the architecture shown in figure 1.

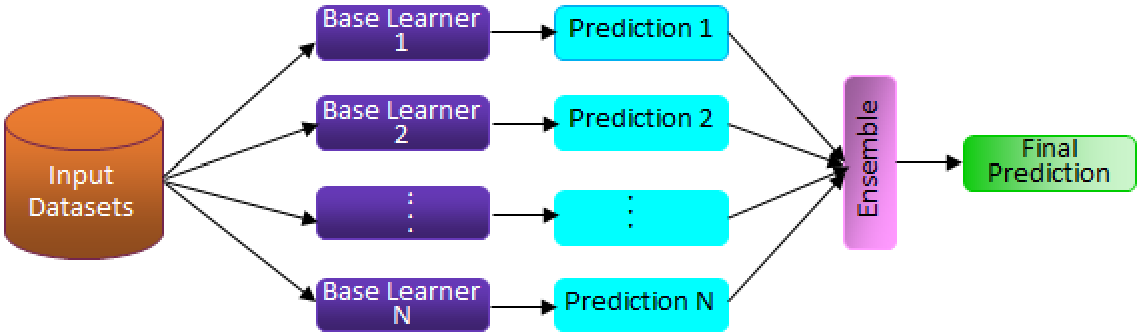

Figure 1: Stacking Ensemble Model Architecture

An inclusive machine learning pipeline for diagnosing heart disease was implemented. Hyperparameter tuning via grid search technique, with cross-validation (CV=3), was implored to identify the optimal set of hyperparameters for each model. For the XGBoost model, different parameters were tested in different combinations to find the most effective configuration as n_estimators: [100, 250], learning_rate: [0.01, 0.2], max_depth: [3, 8], min_child_weight: [1, 3, 5], subsample: [0.6, 1.0], and colsample_bytree: [0.6, 1.0]. The Random forest had the best combination of parameters as n_estimators: [100, 250], max_depth: [6, 12], min_samples_split: [2, 7], and min_samples_leaf: [1, 4]. The support vector machine (SVM) had its parameters set to C: [0.01, 10], gamma: [0.001, 1], and kernel: ['linear', 'rbf'] and the Logistic regression parameters were set to C: [0.01, 10], solver: ['liblinear', 'lbfgs', 'saga'], max_iter : [100, 300]. Next, the stacking ensemble model's effectiveness was assessed. The final model was evaluated using metrics which includes precision, recall, specificity, and the F1 score. The performance of the model was visualized through the confusion matrix and ROC curve. Python modules like Pandas and NumPy were used for data manipulation and preprocessing across the whole model construction, training, and evaluation phase. Logistic regression was used for the Meta classifier, XGBoost, Random Forest, and Support Vector Machine (SVM) were used as base classifiers, and Matplotlib and Seaborn are used for data visualization and analysis. Scikit-Learn is used for model construction and assessment. The use of multiple models and the stacking ensemble approach aimed at improving predictive performance, while rigorous evaluation ensured the reliability of the results.

4. Results and Discussion

The stacking ensemble model was implemented to improve the robustness of heart disease diagnosis using a combination of different models. This section presents the results of the implementation and discusses the performance of the model based on various evaluation metrics and visualizations.

The performance of the stacking ensemble model was evaluated using several key metrics which includes accuracy, precision, recall, F1 score, and specificity. The result yields an accuracy of 89%, Precision of 89%, Recall (sensitivity) of 85%, F1 Score of 87% and Specificity of 92%. These metrics indicate that the stacking ensemble model performs well in diagnosing heart disease, with high precision and recall values, suggesting a good balance between the model's ability to identify positive cases and minimize false positives. A normalized confusion matrix was plotted to provide better insights into the model's performance by showing the proportion of correct and incorrect predictions ad depicted in figure 2.

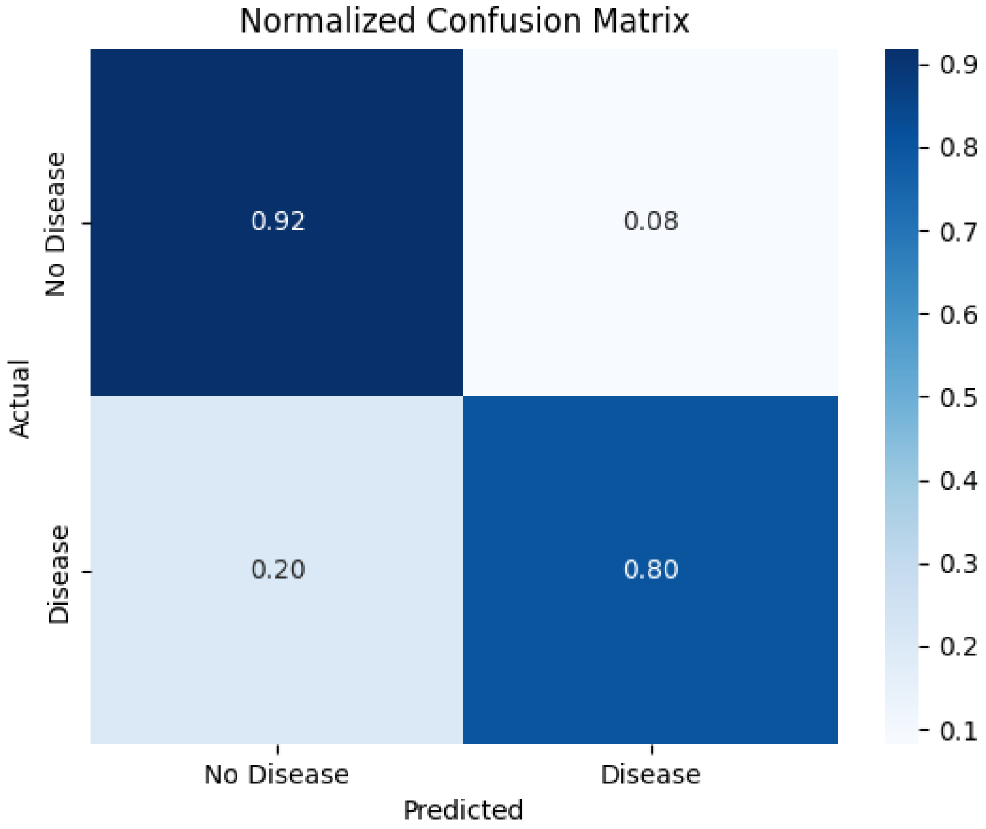

Figure 2: Confusion Matrix plot

The confusion matrix confirms that the model correctly identifies most of the positive and negative cases, with few misclassifications. The model is highly effective (92%) at correctly identifying patients who do not have heart disease and makes few mistakes (8%) in incorrectly predicting heart disease in patients who do not have it, reducing unnecessary worry and potential over-treatment. The Plot further suggests that the model accurately identifies a large majority of patients (85%) with heart disease, which is crucial for timely and appropriate treatment as depicted. The high true positive and true negative rates indicate that the model is generally reliable.

To evaluate the model's discriminatory power, the ROC curve was also plotted and the area under the curve (AUC) computed. The outcome is depicted in Figure 3.

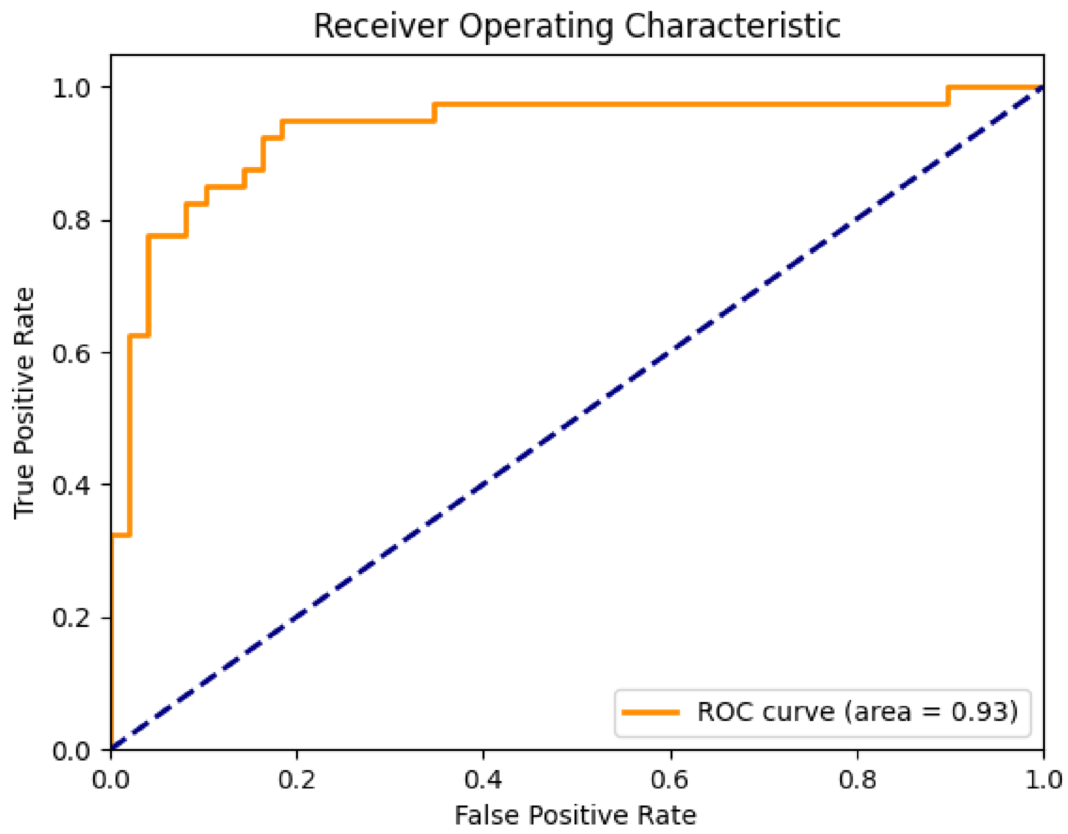

Figure 3: Receiver operating characteristic (ROC) curve

The stacking ensemble model's outstanding discriminative capacity to differentiate between individuals with and without heart disease is indicated by the ROC curve, which has an AUC of 0.93. The model's high AUC shows that it can continue to function well even in the event that the patient population or the underlying data distribution changes. Healthcare providers can be certain that the model works effectively over a range of decision thresholds thanks to the ROC curve's shape and the AUC number.

A modest difference exists between the cross-validation and training scores, as illustrated in the learning curves created for the stacking ensemble model in Figure 4.

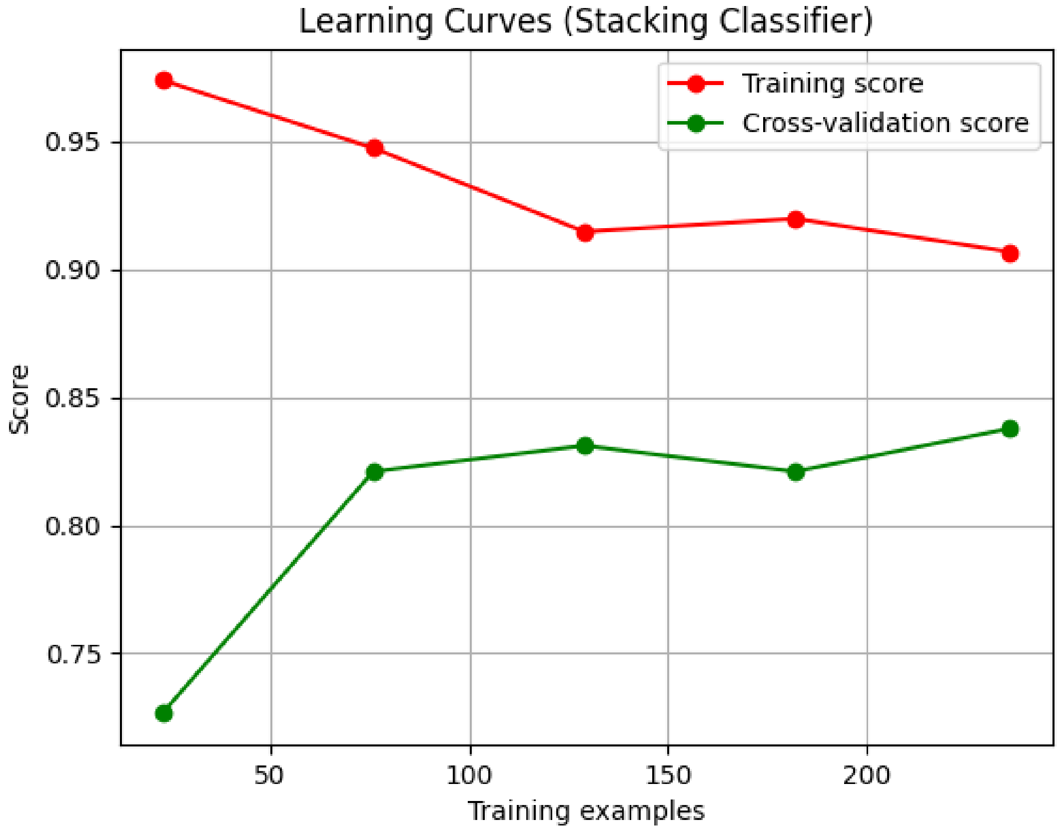

Figure 4: Learning curves

The high cross-validation score and the high training score in this instance show that the ensemble model does not under or over fit the training data. The model appears to be well-generalized and able to produce precise predictions on fresh, untested data based on this balance. All this indicates is that the model has successfully balanced variance and bias. In clinical applications, like the diagnosis of heart disease, where consistent performance across different datasets (training and testing) is critical for reliable outcomes, the ensemble model's robustness is indicated by the close proximity of the training and cross-validation curves.

5. Conclusion

In comparison to conventional single-algorithm methods, the stacking ensemble model developed in this work offers improved accuracy and reliability, making it a potent tool for diagnosing cardiac disease. The obtained result indicates that the model is highly equipped for real-world clinical deployment in heart disease diagnosis, and it has a good chance of retaining accuracy and dependability. The promise of AI-driven solutions to enhance clinical decision-making and patient outcomes in healthcare is highlighted by the successful integration of several machine learning algorithms into a coherent ensemble model. In order to improve this model's prediction power, future research could examine how to apply it to different medical problems and include more data sources. In order for AI-driven healthcare diagnostics to be more extensively recognized and used in clinical practice, it is imperative that key issues like bias, fairness, and trust be addressed by the suggested paradigm. The result demonstrates that the model can increase the accuracy of heart disease detection, supporting medical practitioners in reaching better decisions. The effectiveness of the model was established in practical situations by means of validation through the use of publically accessible datasets on heart disease. The study's findings shows that the stacking ensemble model performs better than individual models, which makes a strong argument for its application in therapeutic contexts.

References

1. M. Bagheri, M. Bagheritaba, S. Alizadeh, M. S. Parizi, P. Matoufinia, and Y. Luo, “AI-Driven Decision-Making in Healthcare Information Systems: A Comprehensive Review,” Jun. 12, 2024. doi: 10.20944/preprints202406.0790.v1.
2. S. H. Bangash, I. Khan, G. Husnain, M. A. Irfan, and A. Iqbal, “Revolutionizing Healthcare with Smarter AI: In-depth Exploration of Advancements, Challenges, and Future Directions,” *VFAST trans. softw. eng.*, vol. 12, no. 1, pp. 152–168, Mar. 2024, doi: 10.21015/vtse.v12i1.1760.
3. P. Esmaeilzadeh, “Use of AI-based tools for healthcare purposes: a survey study from consumers’ perspectives,” *BMC Med Inform Decis Mak*, vol. 20, no. 1, p. 170, Dec. 2020, doi: 10.1186/s12911-020-01191-1.
4. D. Lee and S. N. Yoon, “Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges,” *IJERPH*, vol. 18, no. 1, p. 271, Jan. 2021, doi: 10.3390/ijerph18010271.
5. J. P. Richardson *et al.*, “Patient apprehensions about the use of artificial intelligence in healthcare,” *npj Digit. Med.*, vol. 4, no. 1, p. 140, Sep. 2021, doi: 10.1038/s41746-021-00509-1.
6. S. U. D. Wani *et al.*, “Utilization of Artificial Intelligence in Disease Prevention: Diagnosis, Treatment, and Implications for the Healthcare Workforce,” *Healthcare*, vol. 10, no. 4, p. 608, Mar. 2022, doi: 10.3390/healthcare10040608.
7. M. Alohali, F. Carton, and Y. O’Connor, “Investigating the antecedents of perceived threats and user resistance to health information technology: a case study of a public hospital,” *Journal of Decision Systems*, vol. 29, no. 1, pp. 27–52, Jan. 2020, doi: 10.1080/12460125.2020.1728988.
8. Y. Chen, C. Stavropoulou, R. Narasinkan, A. Baker, and H. Scarbrough, “Professionals’ responses to the introduction of AI innovations in radiology and their implications for future adoption: a qualitative study,” *BMC Health Serv Res*, vol. 21, no. 1, p. 813, Dec. 2021, doi: 10.1186/s12913-021-06861-y.
9. S. A. Alowais *et al.*, “Revolutionizing healthcare: the role of artificial intelligence in clinical practice,” *BMC Med Educ*, vol. 23, no. 1, p. 689, Sep. 2023, doi: 10.1186/s12909-023-04698-z.
10. “Science Magazine - October 25, 2019 - Dissecting racial bias in an algorithm used to manage the health of populations,” 2019.
11. T. L. Upshaw *et al.*, “Priorities for Artificial Intelligence Applications in Primary Care: A Canadian Deliberative Dialogue with Patients, Providers, and Health System Leaders,” *J Am Board Fam Med*, vol. 36, no. 2, pp. 210–220, Apr. 2023, doi: 10.3122/jabfm.2022.220171R1.
12. J. Zhang and Z. Zhang, “Ethics and governance of trustworthy medical artificial intelligence,” *BMC Med Inform Decis Mak*, vol. 23, no. 1, p. 7, Jan. 2023, doi: 10.1186/s12911-023-02103-9.
13. A. Lundgard, “Measuring justice in machine learning,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, Jan. 2020, pp. 680–680. doi: 10.1145/3351095.3372838.
14. S. Caton and C. Haas, “Fairness in Machine Learning: A Survey,” *ACM Comput. Surv.*, vol. 56, no. 7, pp. 1–38, Jul. 2024, doi: 10.1145/3616865.
15. E. Wall *et al.*, “Trust Junk and Evil Knobs: Calibrating Trust in AI Visualization,” in *2024 IEEE 17th Pacific Visualization Conference (PacificVis)*, Tokyo, Japan: IEEE, Apr. 2024, pp. 22–31. doi: 10.1109/PacificVis60374.2024.00012.

16. C. Gomez, B.-L. Smith, A. Zayas, M. Unberath, and T. Canares, “Explainable AI decision support improves accuracy during telehealth strep throat screening,” *Commun Med*, vol. 4, no. 1, p. 149, Jul. 2024, doi: 10.1038/s43856-024-00568-x.

17. C. Metta, A. Beretta, R. Pellungrini, S. Rinzivillo, and F. Giannotti, “Towards Transparent Healthcare: Advancing Local Explanation Methods in Explainable Artificial Intelligence,” *Bioengineering*, vol. 11, no. 4, p. 369, Apr. 2024, doi: 10.3390/bioengineering11040369.

18. V. Sivaraman, L. A. Bukowski, J. Levin, J. M. Kahn, and A. Perer, “Ignore, Trust, or Negotiate: Understanding Clinician Acceptance of AI-Based Treatment Recommendations in Health Care,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–18. doi: 10.1145/3544548.3581075.

19. S. Tonekaboni, S. Joshi, M. D. McCradden, and A. Goldenberg, “What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use”.

20. A. Y. Hannun *et al.*, “Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network,” *Nat Med*, vol. 25, no. 1, pp. 65–69, Jan. 2019, doi: 10.1038/s41591-018-0268-3.

21. Y. Jin *et al.*, “Cardiologist-level interpretable knowledge-fused deep neural network for automatic arrhythmia diagnosis,” *Commun Med*, vol. 4, no. 1, p. 31, Feb. 2024, doi: 10.1038/s43856-024-00464-4.

22. A. H. Ribeiro *et al.*, “Automatic diagnosis of the 12-lead ECG using a deep neural network,” *Nat Commun*, vol. 11, no. 1, p. 1760, Apr. 2020, doi: 10.1038/s41467-020-15432-4.

23. S. W. Smith *et al.*, “A deep neural network learning algorithm outperforms a conventional algorithm for emergency department electrocardiogram interpretation,” *Journal of Electrocardiology*, vol. 52, pp. 88–95, Jan. 2019, doi: 10.1016/j.jelectrocard.2018.11.013.

24. D. B. Olawade, N. Aderinto, G. Olatunji, E. Kokori, A. C. David-Olawade, and M. Hadi, “Advancements and applications of Artificial Intelligence in cardiology: Current trends and future prospects,” *Journal of Medicine, Surgery, and Public Health*, vol. 3, p. 100109, Aug. 2024, doi: 10.1016/j.glmedi.2024.100109.

25. S. J. Patel *et al.*, “Advancements in Artificial Intelligence for Precision Diagnosis and Treatment of Myocardial Infarction: A Comprehensive Review of Clinical Trials and Randomized Controlled Trials,” *Cureus*, May 2024, doi: 10.7759/cureus.60119.

26. A. J. Russak *et al.*, “Machine Learning in Cardiology—Ensuring Clinical Impact Lives Up to the Hype,” *J Cardiovasc Pharmacol Ther*, vol. 25, no. 5, pp. 379–390, Sep. 2020, doi: 10.1177/1074248420928651.

27. P. Bidwai, S. Gite, K. Pahuja, and K. Kotecha, “A Systematic Literature Review on Diabetic Retinopathy Using an Artificial Intelligence Approach,” *BDCC*, vol. 6, no. 4, p. 152, Dec. 2022, doi: 10.3390/bdcc6040152.

28. X. Qian *et al.*, “The effectiveness of artificial intelligence-based automated grading and training system in education of manual detection of diabetic retinopathy,” *Front. Public Health*, vol. 10, p. 1025271, Nov. 2022, doi: 10.3389/fpubh.2022.1025271.

29. M. A. Urina-Triana *et al.*, “Machine Learning and AI Approaches for Analyzing Diabetic and Hypertensive Retinopathy in Ocular Images: A Literature Review,” *IEEE Access*, vol. 12, pp. 54590–54607, 2024, doi: 10.1109/ACCESS.2024.3378277.

30. A. H. Alharbi and H. A. Hosni Mahmoud, “Pneumonia Transfer Learning Deep Learning Model from Segmented X-rays,” *Healthcare*, vol. 10, no. 6, p. 987, May 2022, doi: 10.3390/healthcare10060987.

31. N. M. Elshennawy and D. M. Ibrahim, “Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images,” *Diagnostics*, vol. 10, no. 9, p. 649, Aug. 2020, doi: 10.3390/diagnostics10090649.

32. W. Khan, N. Zaki, and L. Ali, “Intelligent Pneumonia Identification From Chest X-Rays: A Systematic Literature Review,” *IEEE Access*, vol. 9, pp. 51747–51771, 2021, doi: 10.1109/ACCESS.2021.3069937.

33. Z. Wang *et al.*, “Automatically discriminating and localizing COVID-19 from community-acquired pneumonia on chest X-rays,” *Pattern Recognition*, vol. 110, p. 107613, Feb. 2021, doi: 10.1016/j.patcog.2020.107613.

34. I. Rojek, P. Kotlarz, M. Kozielski, M. Jagodziński, and Z. Królikowski, “Development of AI-Based Prediction of Heart Attack Risk as an Element of Preventive Medicine,” *Electronics*, vol. 13, no. 2, p. 272, Jan. 2024, doi: 10.3390/electronics13020272.

35. P. N. Srinivasu, U. Sirisha, K. Sandeep, S. P. Praveen, L. P. Maguluri, and T. Bikku, “An Interpretable Approach with Explainable AI for Heart Stroke Prediction,” 2024.

36. M. Barrett *et al.*, “Artificial intelligence supported patient self-care in chronic heart failure: a paradigm shift from reactive to predictive, preventive and personalised care,” *EPMA Journal*, vol. 10, no. 4, pp. 445–464, Dec. 2019, doi: 10.1007/s13167-019-00188-9.

37. K. Bayoumy *et al.*, “Smart wearable devices in cardiovascular care: where we are and how to move forward,” *Nat Rev Cardiol*, vol. 18, no. 8, pp. 581–599, Aug. 2021, doi: 10.1038/s41569-021-00522-7.

38. G. A. Fleming, J. R. Petrie, R. M. Bergenstal, R. W. Holl, A. L. Peters, and L. Heinemann, “Diabetes Digital App Technology: Benefits, Challenges, and Recommendations. A Consensus Report by the European Association for the Study of Diabetes (EASD) and the American Diabetes Association (ADA) Diabetes Technology Working Group,” *Diabetes Care*, vol. 43, no. 1, pp. 250–260, Jan. 2020, doi: 10.2337/dci19-0062.

39. R. Genuer, J.-M. Poggi, C. Tuleau-Malot, and N. Villa-Vialaneix, “Random Forests for Big Data,” *Big Data Research*, vol. 9, pp. 28–46, Sep. 2017, doi: 10.1016/j.bdr.2017.07.003.

40. D. P. Mohandoss, Y. Shi, and K. Suo, “Outlier Prediction Using Random Forest Classifier,” in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, NV, USA: IEEE, Jan. 2021, pp. 0027–0033. doi: 10.1109/CCWC51732.2021.9376077.

41. T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

42. O. Sagi and L. Rokach, “Approximating XGBoost with an interpretable decision tree,” *Information Sciences*, vol. 572, pp. 522–542, Sep. 2021, doi: 10.1016/j.ins.2021.05.055.