Evolution of Explainable AI in Healthcare: Toward Trustworthy and Accurate Diagnostics

Ibtasam Ur Rehman*, Abdulraqeb Alhammadi†, Jibran K. Yousafzai‡
*Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam
†Faculty of Artificial Intelligence, University of Technology Malaysia
‡College of Engineering and Applied Sciences, American University of Kuwait

***Abstract*—The integration of Artificial Intelligence in healthcare is revolutionizing medical diagnostics, treatment personalization, and clinical decision-making. This survey examines recent advancements in AI-driven healthcare applications, focusing on diagnostic accuracy and algorithmic transparency. Our analysis of research demonstrates that ensemble machine learning models and deep learning approaches are achieving remarkable accuracy rates, often exceeding 95% across various medical domains, including cardiology, neurology, dermatology, and ophthalmology. A significant trend identified is the critical integration of Explainable AI techniques such as SHAP and LIME to address the black-box problem, thereby enhancing clinical trust and adoption. The survey also highlights emerging challenges, including data privacy concerns, algorithmic bias, and the need for regulatory frameworks. By synthesizing findings from recent literature, this paper provides an overview of the current landscape of AI in healthcare, identifying promising research directions and balancing technological innovation with ethical responsibility in medical AI systems.**

***Index Terms*—Artificial Intelligence, Healthcare Diagnostics, Explainable AI, Machine Learning, Deep Learning, Clinical Decision Support, Medical Imaging, Personalized Medicine, Healthcare Ethics, Algorithmic Transparency**

I. Introduction

A. Motivation: Why AI in Healthcare?

The integration of Artificial Intelligence (AI) in healthcare represents a paradigm shift that can address long-standing challenges such as delayed disease detection, generalized treatment strategies, and healthcare inequities. Traditional diagnostic practices often depend on the symptomatic presentation of diseases at advanced stages, while the "one-size-fits-all" treatment model overlooks inter-patient variability. AI introduces data-driven precision through pattern recognition, enabling early detection, individualized treatment planning, and expanded access to care via automated systems. The COVID-19 pandemic further accelerated AI adoption, demonstrating its critical role in rapid diagnostics, predictive modeling, and remote patient monitoring during public-health emergencies [1], [2].

B. Challenges in Current Healthcare

Despite technological progress, healthcare systems continue to face limitations that AI seeks to overcome. Late diagnosis of life-threatening conditions such as Alzheimer's and cardiovascular diseases restricts effective intervention [3], [4]. Generic treatment protocols ignore genetic and lifestyle differences, often resulting in inconsistent outcomes. Moreover, inequities in access and quality persist across geographic and socioeconomic boundaries. Diagnostic inconsistencies arising from human fatigue or cognitive bias further contribute to preventable errors.

C. Scope of This Paper

This paper systematically reviews recent advancements in AI-driven healthcare applications, emphasizing the dual objectives of diagnostic accuracy and algorithmic transparency. It synthesizes and compares studies that implement Machine Learning (ML), Deep Learning (DL), and Explainable AI (XAI) methods, including SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and Gradient-weighted Class Activation Mapping (Grad-CAM), across medical specialties such as cardiology, neurology, dermatology, and ophthalmology. The survey identifies performance trends, highlights clinical use cases surpassing 95% accuracy, and examines how interpretability methods mitigate the black-box challenge. By integrating findings from diverse datasets and methodologies, this review provides a comprehensive perspective on how Explainable AI is reshaping clinical trust, adoption, and reliability in modern healthcare diagnostics.

The main contributions of this paper are as follows:

- A systematic synthesis of XAI techniques applied in diagnostic modeling across multiple clinical domains.
- Quantitative comparison of AI model accuracy, including ML and DL approaches with interpretability integration. The study benchmarks performance differences to reveal how integrating explainable mechanisms influences diagnostic reliability and generalization.
- Identification of challenges and future research trends in deploying transparent AI systems for healthcare. This analysis highlights critical gaps, such as data heterogeneity and interpretability-performance trade-offs, offering directions for future advancements in trustworthy medical AI.

The remainder of this paper is organized as follows. Section II outlines the background of AI in healthcare, covering ML, DL, and XAI fundamentals. Section III presents a comprehensive literature review and comparative analysis summarized in Table I. Section IV discusses clinical applications and diagnostic implementation. Finally, Section V concludes with key insights and future research directions in AI-driven healthcare diagnostics.

II. Background

A. Overview of AI in Healthcare

According to Alhejaily *et al.* [5], AI is transforming modern healthcare by enhancing the precision, accessibility, and efficiency of medical services, particularly in diagnostics, where algorithmic accuracy increasingly parallels that of human experts. While the widespread digitization of health data enables these advancements, the authors emphasize ongoing challenges in data privacy, ethical compliance, and regulatory oversight. Complementing this perspective, Shaheen [6] underscores AI's capacity to mitigate systemic healthcare issues such as rising costs and workforce shortages through innovations in drug discovery, clinical trial optimization, and patient care automation.

Figure 1 summarizes the multifaceted role of AI across the medical ecosystem. The diagram highlights key areas of impact including disease prediction, enhanced diagnostic accuracy through automated image analysis, expanded access to healthcare services and biometric identity verification. Further applications encompass robot-assisted surgeries, personalized treatment for rare diseases, and the deployment of conversational chatbots for patient interaction. Each spoke of the figure represents a critical application domain where AI technologies contribute to a more efficient, precise, and inclusive healthcare system.

Fig. 1: Illustration of AI systems and applications in healthcare.

B. Machine Learning, Deep Learning, and Explainable AI in Healthcare

ML and DL methods are driving major advances in healthcare by expediting drug discovery, optimizing clinical trials, and enhancing diagnostic precision, especially in medical image analysis. Nevertheless, the black-box nature of complex DL architectures often hinders clinical adoption, as medical practitioners require transparency and interpretability to trust algorithmic decisions. To address this challenge, XAI has emerged as a crucial paradigm.

XAI techniques provide interpretability through several complementary approaches. **SHAP** attributes each feature's contribution to individual predictions using cooperative game theory; **LIME** constructs locally faithful approximations by perturbing input data and observing output changes; and **Grad-CAM** generates visual heatmaps that highlight salient image regions influencing a model's decision. Collectively, these techniques ensure that AI systems rely on clinically relevant patterns rather than spurious correlations.

The integration of ML, DL, and XAI is therefore fundamental to developing efficient, transparent, and clinically trustworthy AI-driven tools for patient care [7], [8]. Figure 2 outlines a generalized workflow for designing and refining such systems. The process begins with data collection, preprocessing, and feature engineering, followed by model training, selection (e.g., Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), or Transformer), and rigorous validation. The iterative loop continues until satisfactory performance is achieved, after which the model is deployed within clinical workflows as a decision-support system. Continuous post-deployment monitoring, clinician feedback, and periodic retraining ensure sustained model reliability and relevance in dynamic healthcare environments.

Having established the foundational AI methods and their interpretability mechanisms, the next section presents a structured literature review of their applications in clinical diagnostics.

III. Literature Review

This systematic review employed a structured search and selection process across major databases, including IEEE Xplore, PubMed, and ScienceDirect, covering the period 2020–2025. Search keywords included "AI in healthcare," "Explainable AI," "diagnostic accuracy," and domain-specific terms such as "cardiology" and "neurology." Studies were included if they reported quantitative performance metrics and incorporated XAI methods. The following paragraphs summarize representative studies that integrate explainability within diagnostic AI models across diverse medical domains.

Several studies demonstrate the effective use of Explainable AI (XAI) for building accurate and transparent heart disease prediction models. Sethi *et al.* [4] developed a model focused on medical transparency, achieving 96.07% accuracy by leveraging key clinical features such as peak exercise ST segment slope and maximum heart rate. Within the domain of electrocardiogram (ECG) analysis, Aggarwal *et al.* [9] implemented an XAI-driven system that used morphological characteristics of heartbeat waveforms to classify signals, with Logistic Regression (LR) and Support Vector Machine (SVM) models achieving up to 95% accuracy. Further advancing this

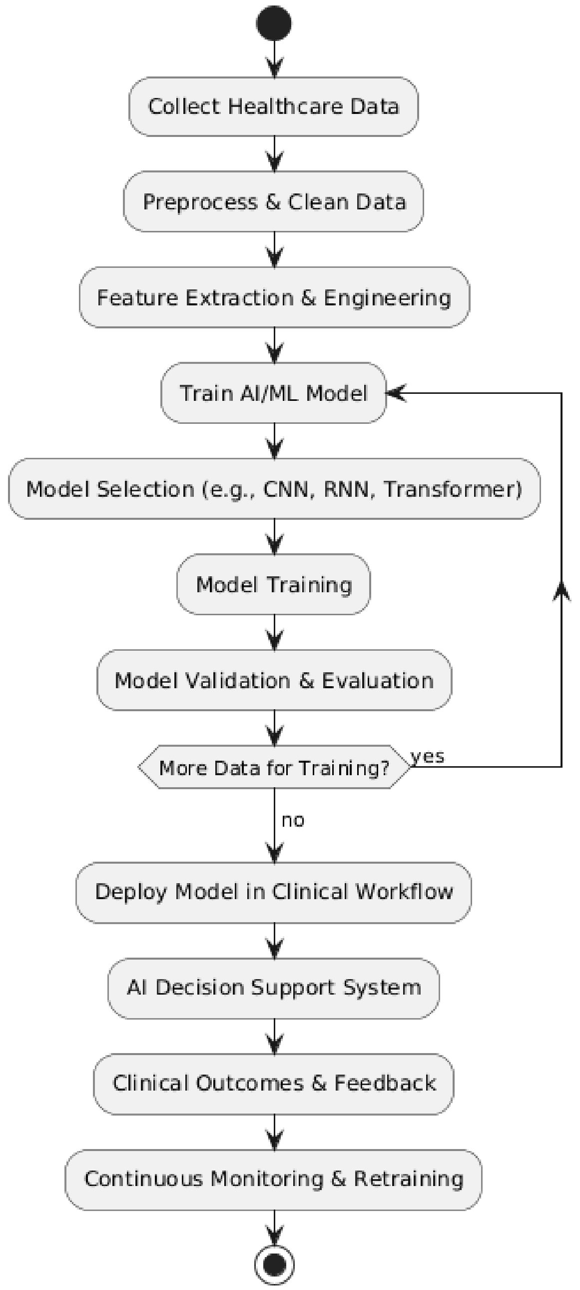

Fig. 2: Generalized workflow for implementing and continuously improving AI systems in healthcare.

field, Rezk *et al.* [10] built a highly accurate Voting Ensemble model (96.5%) combining Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) algorithms. Their work identified cholesterol as the most critical predictor and extensively used SHAP and LIME to provide feature-level explanations for clinical users.

The expansion of transparent diagnostic tools is pivotal for clinical AI integration. Sanchula *et al.* [11] rated multiple models for heart disease prediction using SVM, achieving 95.34% accuracy. To ensure transparency, they used SHAP and LIME to generate both global and local explanations for model predictions. In a related study, a hybrid multi-disease framework using XGBoost achieved 99.2% accuracy and applied LIME to highlight the specific health features, such as cholesterol levels, responsible for each diagnosis, fostering greater trust among healthcare professionals. Guleria *et al.* [12] developed an XAI framework for cardiovascular disease prediction using the Cleveland heart disease dataset (303 patients, 14 features) from the University of California Irvine (UCI) repository. They tested several algorithms, with SVM, LR, and Naive Bayes (NB) achieving the highest accuracy of 89%, while SHAP and LIME explanations enhanced interpretability for clinicians.

Gupta *et al.* [13] evaluated multiple ML algorithms for cardiovascular disease prediction, identifying LR as the most effective with 83.90% accuracy. Using clinical features including age, chest pain type, and blood pressure, their study incorporated XAI techniques to improve model transparency for decision-making. Shijin *et al.* [3] created a multimodal transformer model for predicting the conversion from Mild Cognitive Impairment to Alzheimer's Disease. Their framework, which fuses magnetic resonance imaging (MRI) scans and clinical data, achieved accuracies of 79.76% on ADNI-1 and 86.89% on ADNI-2. In cardiovascular research, Bilal *et al.* [14] developed an XAI-based system using clinical and lifestyle data from over 300,000 patients, achieving 91.94% accuracy with models like LightGBM and XGBoost and used SHAP and LIME for transparency. Yaseen and Rashid [15] found XGBoost to be the most effective model for heart disease classification (92% accuracy, 0.93 Area Under the Curve–Receiver Operating Characteristic (AUC-ROC)), identifying chest pain type and ST depression as key predictors via SHAP analysis. Finally, Rehman and Pham [2] introduced the *Cortex Vision* mobile application, which uses an SVM model with 95% accuracy to detect cataracts from ocular images, demonstrating the potential for accessible AI-driven diagnostics.

Abbas *et al.* [16] developed an intelligent skin disease prediction system using a transfer learning approach with a pre-trained Visual Geometry Group 16-layer (VGG16) network, achieving 93.29% accuracy. To ensure transparency, they incorporated XAI via Layer-wise Relevance Propagation, visually highlighting the rashes and lesions used for diagnosis. In neurological disease prediction, Saimon *et al.* [17] found Gradient Boosting to be most effective for early Parkinson's detection (89% accuracy), while K-Nearest Neighbors achieved 85% accuracy for epilepsy classification from electroencephalogram (EEG) data. For coronary artery disease, Olawade *et al.* [18] employed the Bald Eagle Search optimization method for feature selection. Their Random Forest (RF) model achieved 92% accuracy, with predictors including typical chest pain, ST elevation, and ejection fraction. Esan [19] developed an RF model that achieved 93% accuracy in predicting Parkinson's Disease, with SHAP providing interpretability for clinicians. Similarly, Padhy *et al.* [20] proposed the Weighted Ensemble Explainable AI (WE-XAI) framework for cardiovascular disease prediction, combining high performance with transparency. Abbas *et al.* [21] demonstrated XAI-infused ensemble models using SVM, Decision Tree (DT), and RF, achieving 99% accuracy for heart disease diagnosis. In a different approach, Rezk *et al.* [22] addressed Chronic Kidney Disease prediction using a Generative Adversarial Network (GAN) for data handling and Few-Shot Learning, achieving 99.99% accuracy while maintaining interpretability with SHAP and LIME for clinical trust.

Chowdhury *et al.* [23] developed a hybrid ensemble model for liver disease identification, achieving 98.38% accuracy and applying SHAP and LIME for clinical interpretability. Similarly, Mamun *et al.* [24] introduced a Tree Selection and Stacking Ensemble-based RF model that achieved 99.92% accuracy for liver disease diagnosis, validating key biomarkers aligned with clinical findings. In cardiovascular disease prediction, Kiran *et al.* [25] presented a high-accuracy (98%) RF framework with both global SHAP and local LIME interpretability, enabling clinicians to understand both population-level trends and patient-specific predictions. Li *et al.* [26] developed a clinical prediction model to identify neonates with sepsis who are at risk of developing Purulent Meningitis. Using data from 535 septic neonates, their logistic regression model integrated five key predictors—fever, seizures, tachycardia, and reduced levels of Alkaline Phosphatase and Total Bilirubin—and achieved an AUC of 0.765, offering clinicians an early risk assessment tool. Lin *et al.* [27] employed Least Absolute Shrinkage and Selection Operator (LASSO) regression to differentiate cryptococcal from tuberculous meningitis in Central Nervous System (CNS) infections, achieving AUC 0.919–0.921 in training and validation cohorts, providing a robust diagnostic aid for clinicians.

Chen *et al.* [28] developed a machine learning model to predict short-term adverse outcomes in neonatal bacterial meningitis using data from 433 full-term neonates and 32 clinical variables. Among nine algorithms, LR achieved the best performance with 89.0% accuracy and an AUC of 0.908, identifying muscle tone abnormalities, seizures, and elevated cerebrospinal fluid (CSF) protein as key predictors. Similarly, Sun *et al.* [29] proposed a logistic regression-based nomogram to predict purulent meningitis in 201 very preterm infants (gestational age < 32 weeks), identifying low birth weight, elevated procalcitonin, cesarean delivery, and premature membrane rupture as significant predictors, achieving a concordance index (C-index) of 0.849. Abrar *et al.* [30] developed a CNN-based model for liver disease classification using the Indian Liver Patient Dataset. The CNN model achieved 96.21% accuracy, outperforming other ML models such as SVM and LR, demonstrating deep learning's superior capability for complex medical pattern recognition. This performance underscores CNN's potential as a core component of clinical decision support in hepatology.

The reviewed literature spans cardiovascular, neurological, hepatic, renal, and ophthalmic domains, underscoring the breadth of XAI-enabled diagnostics. As shown in Table I, AI-driven diagnostic systems achieved accuracies ranging from 83% to over 99%, with ensemble and hybrid models generally outperforming single classifiers. Cardiovascular and hepatic disease studies demonstrated the highest accuracies, while neurological applications exhibited slightly lower but still clinically significant performance levels. The reviewed studies collectively illustrate that integrating accuracy with interpretability enhances clinical trust. These findings are further explored through specific clinical applications in Section IV.

IV. AI FOR CLINICAL APPLICATIONS AND DIAGNOSTICS

A. Early Disease Detection and Medical Imaging

AI has revolutionized medical imaging through deep learning architectures such as CNNs and Vision Transformers, enabling automated interpretation of X-rays, CT scans, and MRI data. These systems consistently achieve expert-level accuracy in detecting conditions across multiple specialties, including diabetic retinopathy in ophthalmology, melanoma in dermatology, and cardiac abnormalities in cardiology. XAI methods, particularly Grad-CAM, provide visual interpretability by highlighting the salient image regions that influence diagnostic predictions. Such transparency is essential for verifying model reliability and strengthening clinical confidence in AI-assisted decision-making.

B. Personalized Treatment and Chronic Disease Management

AI facilitates precision medicine by integrating multimodal data sources such as electronic health records, genomic sequences, medical imaging, and wearable-device data to generate individualized treatment recommendations. Machine learning models can predict disease progression and optimize therapeutic strategies for chronic conditions such as diabetes, hypertension, and cardiovascular disorders. The incorporation of continuous real-time physiological data from wearable sensors enables adaptive treatment adjustment, transforming healthcare delivery from a reactive model to a proactive and preventive paradigm that prioritizes early intervention.

C. Mobile Health Applications and Real-World Deployment

AI-powered mobile health applications extend diagnostic and monitoring capabilities directly to patients. Examples include *Cortex Vision* [2] for cataract detection and *DermaAI* for skin-lesion analysis, both demonstrating the potential of smartphone-based diagnostic tools to bridge accessibility gaps in low-resource settings. However, these applications still face notable deployment challenges related to UI/UX design, limited connectivity, device heterogeneity, and disparities in digital literacy. Overcoming these barriers will require not only high algorithmic accuracy but also usability optimization, patient education, and integration within established clinical workflows to ensure sustainable adoption.

TABLE I: Concise Comparative Analysis of AI/ML Approaches in Healthcare Diagnostics

Study	Focus	Model	Features	XAI	Dataset	Contribution	Accuracy
[4]	Cardio	Ensemble ML	ST Segment, HR, Angina	SHAP, LIME	Heart Disease	Transparent Prediction	96%
[12]	Cardio	SVM, LR, NB	Age, Sex, Chol	SHAP	Cleveland	XAI Framework	89%
[9]	Cardio	LR, SVM	ECG Morphology	XAI	ECG Signals	ECG Classification	95%
[10]	Cardio	XGB, LGBM	Chol, BP, Chest Pain	SHAP, LIME	Clinical Data	Voting Ensemble	96%
[11]	Cardio	SVM	Age, Chol, ST-Depression	SHAP, LIME	Heart Disease	Interpretable ML	95%
[13]	Cardio	Logistic Reg	Chest pain, BP, glucose	XAI	Clinical data	Model comparison	83%
[3]	Neuro	Transformer	MRI, Clinical Assessments	-	ADNI	Multimodal fusion	86%
[14]	Cardio	LGBM, XGB	BMI, lifestyle factors	SHAP, LIME	308K Records	Large-scale prediction	91%
[15]	Cardio	XGB	Chest pain, STD	SHAP, LIME	Clinical data	Feature importance	92%
[2]	Ophthal	SVM-RBF	Texture features	-	Ocular images	Mobile application	95%
[16]	Derm	VGG16	Skin lesions	LRP	Skin images	Transfer learning	93%
[17]	Neuro	GBM	Clinical, EEG data	-	Neuro data	Early detection	85%
[18]	Cardio	RF	Chest pain, ST Elev, CR, EF	-	Framingham	BES Feature Selection	92%
[19]	Neuro	RF	UPDRS, MoCA	SHAP, LIME	Clinical data	PD Prediction	93%
[20]	Cardio	Ensemble	Weighted features	SHAP	Clinical data	WE-XAI framework	High
[21]	Cardio	SVM, DT, RF	Clinical attributes	XAI	303 instances	Ensemble + XAI	99%
[22]	Renal	ProtoNets	Clinical markers	SHAP, LIME	CKD	GAN + Few-shot	99%
[23]	Hepatic	ANN hybrids	Bilirubin, Enzymes	LIME, SHAP	Liver disease	Ensemble voting	98%
[24]	Hepatic	TSRF	Biomarkers	SHAP, LIME	Liver disease	Tree selection	99%
[25]	Cardio	RF	Clinical features	SHAP, LIME	Clinical data	Global+local XAI	98%
[26]	Neonatal	Logistic Reg	Fev, Seiz, HR, ALP, TBIL	-	535 neonates	Clinical risk model	AUC: 0.765
[27]	Neuro	LR (LASSO)	CSF pressure, IIS-EF	SHAP	Clinical data	Dx nomogram	AUC: 0.92
[28]	NBM	Multiple	Seizures, HI, CSF	-	433 Neonates	Data driven Prediction tool	98%
[30]	Liver	CNN	Clinical Lab Data	-	ILPD	Enhanced Diagnostic Power	96.21%

Abbreviations: XGB (XGBoost), LGBM (LightGBM), SVM (Support Vector Machine), LR (Logistic Regression), NB (Naive Bayes), RF (Random Forest), DT (Decision Tree), GBM (Gradient Boosting Machine), ANN (Artificial Neural Network), TSRF (Tree Selection Random Forest), ProtoNets (Prototypical Networks), Cardio (Cardiovascular), Neuro (Neurological), Derm (Dermatology), Ophthal (Ophthalmology), Hepatic (Liver), Renal (Kidney), Neonatal (Neonatal Sepsis Meningitis), BP (Blood Pressure), HR (Heart Rate), ECG (Electrocardiogram), MRI (Magnetic Resonance Imaging), EEG (Electroencephalogram), CP (Chest Pain), ST Elev (ST Elevation), Cr (Creatinine), EF (Ejection Fraction), ALP (Alkaline Phosphatase), TBIL (Total Bilirubin), UPDRS (Unified Parkinson's Disease Rating Scale), MoCA (Montreal Cognitive Assessment), PD (Parkinson's Disease), CKD (Chronic Kidney Disease), XAI (Explainable AI), SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), LRP (Layer-wise Relevance Propagation), BES (Bald Eagle Search), FS (Feature Selection), GAN (Generative Adversarial Network), Fev (Fever), Seiz (Seizures), IIS-EF (IIS-Extracranial Fungi), NBM (Neonatal Bacterial Meningitis), HI (Hypotension Inotropes), Chol (Cholesterol),

V. Conclusion and Future Work

This survey validates AI's transformative role in healthcare diagnostics, with models achieving over 95% accuracy across multiple medical domains. The incorporation of XAI techniques such as SHAP and LIME has proven essential for enhancing interpretability, building clinical trust, and supporting adoption. Despite these advancements, persistent challenges remain concerning data privacy, algorithmic bias, and regulatory oversight.

Future research should focus on improving model interpretability and generalizability across diverse populations through federated learning approaches. The development of standardized validation frameworks for clinical deployment and robust bias-mitigation strategies is equally critical. Embedding explainability within federated, privacy-preserving learning pipelines, combined with multimodal data fusion and prospective clinical validation, will be key to delivering trustworthy and generalizable AI systems at scale.

Advances in adaptive learning systems and multimodal data integration will further accelerate personalized medicine, while improved human-AI collaboration interfaces can enhance usability and promote clinical acceptance. This review is limited by the heterogeneity of datasets and reporting metrics across studies, which may affect cross-study comparability. Ultimately, sustained interdisciplinary collaboration among researchers, clinicians, and policymakers will determine the ethical, reliable, and widespread integration of AI in real-world medical practice.

References

[1] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature medicine*, vol. 25, no. 1, pp. 44–56, 2019.

[2] I. Rehman and H.-A. Pham, "Cortex vision: Detection of ophthalmic disease using machine learning algorithm," in *International Conference on Smart Objects and Technologies for Social Good*. Springer, 2024, pp. 138–149.

[3] S. K. GU, A. Purushothaman *et al.*, "Alzfusionformer: Integrating multiple transformers for early alzheimer's disease detection from multi-modal data," *Biomedical Signal Processing and Control*, vol. 112, p. 108601, 2026.

[4] A. Sethi, S. Dharmavaram, and S. K. Somasundaram, "Explainable artificial intelligence (xai) approach to heart disease prediction," in *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)*, 2024, pp. 1–6.

[5] A.-M. G. Alhejaily, "Artificial intelligence in healthcare," *Biomedical Reports*, vol. 22, no. 1, pp. 1–8, 2025.

[6] M. Y. Shaheen, "Applications of artificial intelligence (ai) in healthcare: A review," *ScienceOpen Preprints*, 2021.

[7] P. Kandhare, M. Kurlekar, T. Deshpande, and A. Pawar, "A review on revolutionizing healthcare technologies with ai and ml applications in

pharmaceutical sciences," *Drugs and Drug Candidates*, vol. 4, no. 1, p. 9, 2025.

[8] A. Chaddad, Y. Hu, Y. Wu, B. Wen, and R. Kateb, "Generalizable and explainable deep learning for medical image computing: An overview," *Current Opinion in Biomedical Engineering*, vol. 33, p. 100567, 2025.

[9] R. Aggarwal, P. Podder, and A. Khamparia, "Ecg classification and analysis for heart disease prediction using xai-driven machine learning algorithms," in *Biomedical data analysis and processing using explainable (XAI) and responsive artificial intelligence (RAI)*. Springer, 2022, pp. 91–103.

[10] N. G. Rezk, S. Alshathri, A. Sayed, E. El-Din Hemdan, and H. El-Behery, "Xai-augmented voting ensemble models for heart disease prediction: A shap and lime-based approach," *Bioengineering*, vol. 11, no. 10, p. 1016, 2024.

[11] S. A. Sanchula, "Explainable ai (xai) for a machine learning heart disease prediction model," 2025.

[12] P. Guleria, P. Naga Srinivasu, S. Ahmed, N. Almusallam, and F. K. Alarfaj, "Xai framework for cardiovascular disease prediction using classification techniques," *Electronics*, vol. 11, no. 24, p. 4086, 2022.

[13] R. Gupta, A. Panwar, S. Manhas, K. Ali, and M. Saraswat, "Cardiovascular disease prediction using machine learning and xai," in *2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*. IEEE, 2025, pp. 752–758.

[14] D. A. Bilal, A. Alzahrani, K. Almohammadi, M. Saleem, M. S. Farooq, and R. Sarwar, "Explainable ai-driven intelligent system for precision forecasting in cardiovascular disease," *Frontiers in Medicine*, vol. 12, p. 1596335, 2025.

[15] O. M. Yaseen and M. M. Rashid, "An explainable artificial intelligence (xai) methodology for heart disease classification."

[16] S. Abbas, F. Ahmed, W. A. Khan, M. Ahmad, M. A. Khan, and T. M. Ghazal, "Intelligent skin disease prediction system using transfer learning and explainable artificial intelligence," *Scientific Reports*, vol. 15, no. 1, p. 1746, 2025.

[17] S. I. Saimon, I. Islam, S. I. Abir, N. Sultana, M. S. Hossain, and S. A. Al Shiam, "Advancing neurological disease prediction through machine learning techniques," *Journal of Computer Science and Technology Studies*, vol. 7, no. 1, pp. 139–156, 2025.

[18] D. B. Olawade, A. A. Soladoye, B. A. Omodunbi, N. Aderinto, and I. A. Adeyanju, "Comparative analysis of machine learning models for coronary artery disease prediction with optimized feature selection," *International Journal of Cardiology*, p. 133443, 2025.

[19] A. O. Esan, D. B. Olawade, A. A. Soladoye, B. A. Omodunbi, I. A. Adeyanju, and N. Aderinto, "Explainable ai for parkinson's disease prediction: A machine learning approach with interpretable models," *Current Research in Translational Medicine*, p. 103541, 2025.

[20] S. K. Padhy, A. Mohapatra, and S. Patra, "We-xai: explainable ai for cvd prediction using weighted feature selection and ensemble classifiers," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 14, no. 1, p. 13, 2025.

[21] N. Abbas, T. Tasleem, A. Hai, Z. R. Alqahtani, and B. A. M. A. Alghamdi, "Enhancing heart disease diagnosis with xai-infused ensemble classification," in *Explainable Artificial Intelligence in Medical Imaging*. Auerbach Publications, 2025, pp. 147–166.

[22] N. G. Rezk, S. Alshathri, A. Sayed, and E. E.-D. Hemdan, "Explainable ai for chronic kidney disease prediction in medical iot: Integrating gans and few-shot learning," *Bioengineering*, vol. 12, no. 4, p. 356, 2025.

[23] S. H. Chowdhury, M. Mamun, T. A. Shaikat, M. I. Hussain, S. Iqbal, and M. M. Hossain, "An ensemble approach for artificial neural network-based liver disease identification from optimal features through hybrid modeling integrated with advanced explainable ai," *Medinformatics*, vol. 2, no. 2, pp. 107–119, 2025.

[24] M. Mamun, S. H. Chowdhury, M. M. Hossain, M. Khatun, and S. Iqbal, "Explainability enhanced liver disease diagnosis technique using tree selection and stacking ensemble-based random forest model," *Informatics and Health*, vol. 2, no. 1, pp. 17–40, 2025.

[25] I. Kiran, S. Ali, M. Alhussein, S. Aslam, K. Aurangzeb *et al.*, "An ai-enabled framework for transparency and interpretability in cardiovascular disease risk prediction." *Computers, Materials & Continua*, vol. 82, no. 3, 2025.

[26] J. Li, C. Song, T. Li, W. Jia, Z. Qian, Y. Peng, Y. Xu, and Z. Jin, "Construction and validation of a nomogram model for predicting the risk of neonatal sepsis complicated by purulent meningitis," *Journal of Inflammation Research*, pp. 7183–7194, 2025.

[27] F. Lin, C. Chen, Y. Li, and H. Liu, "Machine learning–based diagnosis of cryptococcal meningitis and tuberculous meningitis: A single-center retrospective clinical study," in *Open Forum Infectious Diseases*, vol. 12, no. 5. Oxford University Press US, 2025, p. ofaf217.

[28] Y. Chen, S. Wang, C. Wang, N. Zhang, Y. Li, H. Shi, P. Zou, H. He, and Y. Wang, "Establishment of a machine learning-based prediction model for short-term adverse prognosis in neonatal bacterial meningitis," *iLABMED*, vol. 3, no. 3, pp. 292–302, 2025.

[29] X. Sun, R. Jing, and Y. Li, "Predicting purulent meningitis in very preterm infants: a novel clinical model," *BMC pediatrics*, vol. 25, no. 1, p. 3, 2025.

[30] A. A. Syed, A. B. Shaik, S. Konatam, M. V. P. Kumar, and N. Gupta, "Cnn-based analytical approach for liver disease classification and prediction," in *2025 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*. IEEE, 2025, pp. 682–685.