International Journal of Basic and Applied Sciences, 14 (8) (2025) 240-250

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/mq9yvg65
Research paper

An Efficient Hybrid Model for Healthcare System to Detect Disease Using Machine Learning Techniques

Ajay Singh Mavai [1] *, Devendra Kumar Mishra [2], Abhishek Sharma [3]

[1] PhD Scholar, Amity University Madhya Pradesh, Gwalior, M.P. India
[2] Amity University Madhya Pradesh, Gwalior, M.P. India
[3] The LNM Institute of Information Technology, Jaipur, Rajasthan, India
**Corresponding author E-mail: ajaymavai@gmail.com*

Received: August 25, 2025 Accepted: November 4, 2025, Published: December 12, 2025

Abstract

Over the past decade, the huge demand for personalized and proactive healthcare has led to the integration of advanced technologies such as wearable devices, IoT and AI into modern healthcare. The project proposes the design of a smart health care system specifically for elderly people and disabled persons, which continuously monitors their health and provides early disease prediction. It proposes integrating a system of patient medical records with real-time physiological data collected by multiple IoT-based wearable devices, while ensuring reliable and timely assistance. IoT-based networks work on standardized communication protocols to promote and manage data efficiency. In addition, it uses sophisticated machine learning algorithms to evaluate collected data, identify potential health risks and support clinical decision-making with utmost accuracy and precision. It enhances the overall effectiveness and scalability of healthcare service in addition to improving early treatment outcomes. Thus, by addressing issues of heterogeneity, missing values, and interpretability, this proposed strategy advances the creation of an inclusive healthcare system.

Keywords: *Healthcare; Machine Learning; Personalized Medicine; Predictive Analytics; Prognosis and Disease Prediction.*

1. Introduction

IoT-based wearable technology combined with machine learning algorithms is moving the healthcare sector away from traditional, reactive approaches and towards more proactive, preventive and personalized care. Because they are more likely to experience underdiagnosed conditions and delayed diagnoses as a result of irregular health monitoring, this change is especially critical for the elderly and disabled.
To answer these challenges, this research addresses three objectives: one, the design of a smart healthcare system integrating previous medical data with real-time health data collected using wearable sensors for the disabled and the elderly; second, leveraging IoT-based sensor networks and protocols for secure, scalable, and reliable collection and data transmission; and third, application of AI & ML-driven predictive models for helping with the identification of diseases with high accuracy, prediction of prognosis, and making clinical decisions judiciously.In this paper we especially focus on third objective that emphasize on reliable, accurate and efficient predictive analytics model which the implement Machine Learning and Artificial Intelligence-based intelligent algorithm for Decision making and prediction.
Predictive Analytics is the "Pillar of Smart Healthcare System", which is currently a vital health care tool that can potentially transform patient care, resource planning, and public health management. According to WHO, 1 in 6 people will be 60 years or older in 2030, and predictive technologies will be of highest significance for managing age-related chronic diseases [1]. Patterns that are not apparent in Electronic Health Records (EHRs), genomic data, sensor data, and behavioral data can be identified by ML algorithms to make early disease detection, improve the accuracy of prognosis, and suggest personalized preventive interventions.
Healthcare practitioners increasingly turn to predictive models for: Early Diagnosis of Disease: Studies show that ML models can make disease diagnoses such as diabetes, Parkinson's disease, and cardiac defects with more than 90% accuracy when trained on multimodal data [2]. Personalized Prognosis: A patient's demographic, lifestyle, and genetic data are utilized by prediction models to forecast disease outcome and aid clinicians in developing treatment plans [3]. Preventive Interventions: Risk stratification by analytics allows high-impact screening and lifestyle modification, reducing disease incidence as well as improving public health outcomes [4]. Using IoT and wearable devices to collect data: The development of IoT devices enables real-time, non-invasive health monitoring through implantable devices, smart watches, biosensors and fitness trackers. The devices enable the transmission of high-frequency physiological signals such as movement patterns, body temperature, SPO_2 and heart rate wirelessly through protocols such as HTTP/REST API, MQTT and CoAP for processing [5]. For disabled and elderly patients, wearables provide:

- Autonomy and independence through passive supervision.
- Abnormalities detected at an early stage, warnings given before clinical symptoms worsen.
- Need-based help, for example, falls checks or reminders to take medication [6].

AI and ML for Predictive Modeling and Decision Support: Predictive and decision support models rely on machine learning algorithms to process and interpret large volumes of clinical and sensor data. Methods used include supervised (eg, random forest, SVM), unsupervised clustering, and deep learning (eg, CNN for images, LSTM for time series health signals). The models provide:

- Very high accuracy in detecting diseases like Alzheimer's, depression, arrhythmia and COPD [7].
- A clinical decision support system (CDSS) that is real-time and provides real-time support during the healthcare process [8].

Resource optimization by predicting hospital readmission, emergency department and bed utilization rates [9].

In addition to its promise, predictive healthcare also faces difficulties such as data heterogeneity, class imbalance, interpretability of black-box models, and ethical issues such as privacy, fairness, and algorithmic bias. The present study addresses these challenges by integrating:

- Explainable AI (XAI) architecture to enhance interpretability.
- Input data treatment including imputation and normalization to handle missing or noisy data.
- Process for healthcare information standards such as HIPAA and GDPR compliance [10].

Machine learning (ML) techniques have been increasingly applied in healthcare predictive analytics, revolutionizing disease detection, prognosis and personalized treatment planning. In general, these developments can be traced to earlier diagnosis, improved patient outcomes, efficient use of resources, and strengthening of clinical decisions through evidence-based information.

1.1. Early disease prediction and prognostic estimation

Early detection of disease to promote early intervention. Machine learning (ML) models have been used to predict various diseases with very high precision. For instance, the application of deep learning on images and genetics has benefited the early detection of cancer and proper treatment selection. Detecting high-risk patients before the onset of symptoms, improving treatment outcomes, and increasing survival rates have been shown in the research done on breast, lung, prostate, and colorectal cancers through predictive models [11] [12]. All these activities are aimed at disease burden reduction through preventive management. ML similarly lent its hand in predicting the risk of CVD by looking at the factors of age, blood pressure, cholesterol level, and lifestyle contributions. Doctors can use risk stratification based on the prediction models to assign patient groups to specific prevention and management measures with an individualized treatment plan as the ultimate goal [13]. Spotting people at risk of stroke or heart attack early can be a lifesaver and also a money-saver for the healthcare systems.

1.2. Population health and chronic disease management

The management of chronic conditions like diabetes, hypertension and neurological disorders involves a difficult and long-term process that includes regular monitoring and responsive care. ML methods facilitate the forecasting of disease progression, the forecasting of treatment response, and the forecasting of complications, thus making it possible for healthcare professionals to deliver personalized care plans and home-based patient monitoring services [14] [15]. Predictive data, therefore, supports this solution in reaching its objective of increasing patient compliance and reducing hospital readmissions. Moreover, predictive analytics not only assists population health management by investigating epidemiological and environmental data to reveal disease and healthcare access hotspots and disparities but also supports indirect population health management by providing e.g. machine learning-based models predicting infectious disease transmission such as the COVID-19 pandemic, and thus playing an important role in forecasting outbreak patterns and shaping resource planning accordingly [16]. This capability is in line with the objective of improving public health management through policy and interventions that are not only evidence-based but also targeted.

1.3. Enhancing clinical decision support and healthcare operations

A core focus is on providing robust, interpretable predictive models to aid in clinical decision-making—by utilizing robust machine learning techniques, healthcare providers can analyze large, complex medical datasets (e.g., EHRs, medical imaging studies, genomic data) and provide actionable, real-time clinical information. Machine learning techniques such as feature importance analysis and interpretability frameworks help establish clinician confidence in the predictive capabilities of these models and facilitate the integration of these tooling solutions into standard daily clinical practice [17] [18]. Machine learning-based predictive analytics are also essential to improving and optimizing healthcare operations by generating accurate estimates of patient movement patterns, hospital admission levels and demand for caregiving resources. Improving these predictive abilities will increase the ability to efficiently provide care. Better predictions can reduce patient wait times, create better work distribution among clinicians and staff, and promote more effective utilization of critical resource allocations (ICU beds and ventilators) [19]. Collectively, these operational improvements support long-term viability for healthcare systems while providing consistently high-value to all patients.

1.4. Multimodal data fusion and explainability for clinical trust

The predictive performance of single modality models has been significantly improved over the years by multimodal fusion, which combines real-time streams from electronic health records, imaging data, genomic profiles, and sensors. A major reason for this progress is the acquisition of complementary clinical indications and reference information [27][30]. In terms of the architecture of these fusion methods, they can be classified into three types: initial fusion, followed by input-level concatenation and strict generalization; late fusion, in which predictions are combined by a specific method; and hybrid attention-based architectures, including cross-modal transformers or graph neural networks capable of learning joint representations while preserving the specific structure of each modality. More recent literature reviews emphasize that attentional methods and transformer-based cross-modal encoders, in particular, show greater signing capability for Al-EC. Time-series data with snapshots of episodic laboratory results and imaging to provide more timely and accurate estimation of prognosis [27] [30].To successfully implement multimodal pipelines in a real-world setting, standardization of Preprocessing methods such as unit harmonization and temporal alignment, and ensuring reliable imputation for missing modalities and compliance with FAIR data stewardship principles is crucial. Unless the above critical steps are circumvented, fusion models are at high risk of brittleness and poor generalization across diverse clinical sites.

The successful deployment of clinical practices at the clinical front, however, not only relies on the performance of the involved technologies but also significantly on their interpretability and the extent of trust induced by them. To make the output produced by these models understandable, query-able, and challengeable by the clinicians effectively, it is indispensable to integrate Explainable AI (XAI) techniques into both the model design phase and the deployment phase. Some of the practical XAI approaches are intrinsically interpretable architectures such as sparse additive models and concept-bottleneck networks, and post-hoc explanation tools such as SHAP, Integrated Gradients, and attention maps. These explanation tools are designed for localizing contributing features across modalities and providing counterfactual or case-based explanations that are highly relevant to clinical reasoning processes [28] [29]. It is important to note that the end-user co-design process—cared for by clinicians, nurses, and biomedical engineers—gives rise to specification formats that are appropriate for all types of actions. This could mean, for example, combining symptom-level contributions with clinically-appropriate suggested next steps, making these systems more useful. Additionally, this co-design process allows user-friendliness testing in artificial workflows that are very similar to real settings. It therefore follows that the evaluation of multimodal, explainable systems has to cover not only traditional metrics such as AUROC but also more developed systems [27] [28] [30].

1.4.1. Limitations and challenges of multimodal fusion

Even though the latest attention-based modal and cross-modal transformers offer superior performance, they bring many challenges in their way. Initially, the computational complexity and memory requirements for such architectures are much higher than simple early or late fusion methods, thus, real-time deployment in clinical settings is restricted [31]. Further, accurate alignment of modalities across time and meaning is still a difficult task because clinical data streams typically vary in sampling frequency, duration, or completeness, making the model brittle if any modality is missing or misaligned [32]. Third, interpretability is often lost: as opposed to simple early fusion (simple coupling) or late fusion (combination of unimodal predictions) methods, transformer-based architectures act more like black boxes, raising concerns for clinical validation and regulatory approval [31][33]. In addition, algorithmic bias is a significant risk factor; Previous research has suggested that healthcare AI systems, if not subjected to careful auditing, may inadvertently reproduce or exacerbate existing disparities in the system based on performance assessed for different groups [34]. The aforementioned limitations bring to the fore the need to evaluate each of the mentioned features, such as predictive efficiency, interpretability, computational feasibility, and fairness, before being able to safely incorporate multimodal models into clinical practice.

Various multimodal fusion strategies have been experimented in the healthcare domain with varying pros and cons regarding performance, robustness, and interpretability. Early fusion techniques, where features are combined at the input stage, are still popular because of their simplicity and low computational requirements, however, they have the disadvantage of being easily affected when one of the methods is missing or suffers from misalignment [31]
[32]. Late fusion, on the other hand, is an approach where predictions are derived from models developed for different modalities separately and then combined, this method speaks the same language as the former being more robust and easier to understand for lack of inputs, but still does not take full advantage of the interactions between modalities [31][33]. New hybrids, especially those based on attention and transformers, are achieving predictive accuracy because they are able to learn very complex joint representations and synchronize different types of signals, but they bear the burden of high computational requirements and are less transparent [31]
[33]. The strategy highlights the need for careful selection of approaches based not only on criteria of accuracy but also on feasibility, feasibility and clinical interpretation.

Table 1: Comparison of Multimodal Data Fusion Strategies in Healthcare Applications.

Fusion Type	Description	Advantages	Limitations	Example References
Early Fusion	Concatenates raw or preprocessed features from all modalities at the input layer	Simple to implement, low computational overhead, captures joint interactions directly	Sensitive to scale mismatch, cannot handle missing modalities well, less robust to noise	[31][32]
Late Fusion	Combines independent unimodal predictions using voting, averaging, or meta-learners	Modular, interpretable, robust when modalities are missing, easier to debug	Loses cross-modal feature interactions, may underperform advanced methods	[31][33]
Hybrid / Attention-Based Fusion	Learns joint representations via transformers, attention layers, or graph neural networks	Strong performance, can align heterogeneous data sources, flexible modeling	High computational and memory cost, less interpretable, harder to validate clinically	[31][33]

Although our hybrid fusion approach improves classification metrics, its clinical translation will require addressing practical limitations. In particular, computational and memory demands may limit real-time deployment; data alignment and missing-modality scenarios need robust handling to avoid performance drops in noisy clinical workflows; and fairness audits are essential because label proxy choices can reproduce systemic biases, as previously shown. Future work should include optimization for deployment, domain generalization experiments, and a formal fairness evaluation.

1.5. Privacy-preserving collaboration

Federated Learning and Governance: Federated learning (FL) has been demonstrated to be a very practical and efficient way of training strong prediction models in a large number of healthcare institutions with the assurance that patient information stays localized at its source. Rather than centralizing raw electronic health records (EHRs), imaging, or sensor streams to a universal centralized store, FL explicitly groups model updates—gradients or weights—across all sites and subsequently pools these updates (e.g., using a method called Federated Averaging) to ultimately generate a complete global model. This new mechanism greatly decreases the legal and ethical resistance to sharing data cross-institutionally that is so widespread. Moreover, the method has been applied in radiology, multicenter risk prediction etc., and the generalizability of the model has been enhanced without disclosing any confidential records [26]. To Further enhance privacy, the technique used in FL is often supplemented by a series of technical security measures. Security aggregation protocol that protects the individual updates from being seen by the coordinator and, when needed, advanced cryptographic techniques such as, homomorphic encryption or a trusted execution environment that provide additional security. These multi-layered approaches enable organizations to securely participate in the generation of large training datasets while complying with strict GDPR/HIPAA and organizational governance policies [26], [29].

Nonetheless, it should be pointed out that federated learning (FL) still has its own technical and operational difficulties which are not uncommon but have to be handled very cautiously. One of the main problems revolves around the existence of a non-IID data distribution which is caused by a variety of factors such as demographics, disease prevalence or differences among sites in terms of device features. The mentioned factors might slow the process of federated learning or in case of improper management introduce bias to the global model. The use of strong validation methods is one of the ways to maintain the accuracy and reliability of the results [26], while another option could be the implementation of different algorithmic adaptations, possibly including the establishment of a separate federated model for each hospital site, data reweighting strategies or the application of meta-learning techniques.

The successful implementation of federated learning goes beyond merely having an algorithm that adjusts to input; there needs to be an established governing body or governing entity involved with the project. A governing entity establishes a cooperative agreement among consortium members, creates a data ontology that will be used to define data and its use case via a standardized format, ensures an adequate tracking mechanism is in place to assess provenance, and implements performance/operational monitoring mechanisms (e.g., feedback loops). Without these fundamental characteristics of a governing body, there is a lack of accountability and therefore no clinical safety for patients during the use of federated learning. In addition to establishing the governance structure for FL, to obtain approval from achieving these goals, will involve thorough validation processes (e.g., calibration, decision curve analysis, physician Utility, and clinical validation). These validation processes are crucial because they demonstrate not only the statistical accuracy associated with using FL, but also the significantly enhanced quality of decision-making and patient outcomes that using federated learning can achieve.

Ultimately, developing data ontology must also be part of a prospective multi-site study, including using a Model Card outlining the intended use, uses, limitations, and limitations of the models in this study. In addition to developing data ontology through a prospective multi-site study, you will also need to develop operational guidelines to manage the lifecycle of the data (e.g., developing retraining schedules, develop rollback policies). Also, you will need to develop robust Incident Response Plans for data used. This is important for converting FL prototypes into reliable clinical tools that can be effectively applied in practice [26][29].

1.6. Addressing challenges

Data Quality, Ethics, and Clinical Validation: Clinical AI has made enormous progress but it is not completely developed but so it has to overcome many challenges regarding data quality, moral issues and medical validation before the use of its complete electricity. Healthcare facts is very diverse in nature because it comes from multiple sources together with hospitals, laboratories and various gadgets thus there is no uniformity in elements like statistics layout, completeness and reliability. The presence of those differences ends in poorer overall performance of the fashions and much less applicability to different cases, therefore endangering patients and eroding the trust in the practitioners involved [20][21]. furthermore, the biases that exist inside the ancient data, as an instance the lack of illustration of certain demographic companies, could be taken over through the predictive models leading to biased effects that allows you to harm the affected organizations the most and for that reason inequality. So one can hold predictive electricity when applying in numerous scientific settings, sturdy validation must be finished via potential studies and outside multi-middle reviews [22].

Ethical Considerations: Ethical challenges in fitness care AI cross properly past questions of privacy. despite the fact that techniques inclusive of federated gaining knowledge of [41] are vital for secure, collaborative version education, they do now not with the aid of themselves ameliorate broader moral dangers. Algorithmic bias throughout demographic agencies, which include age, sex, ethnicity, and socioeconomic repute, remains a problem. for instance, studies have proven that predictive models may have decrease performance amongst minority populations, leading to inequitable medical hints and further exacerbating existing health disparities [42][43]. Equitable access to AI-enhanced health care tools is a key dimension because access may vary based on an institution's resources and location, creating two classes of health care institutions—those that can access the tools (i.e., well-resourced, highly trained institutions) and those that are unable to access the tools (i.e., less-resourced, untrained institutions) with limited benefit [44]. To mitigate these inequities, a focused effort to ensure that AI-enhanced health care tools are developed and deployed fairly and equitably requires intentionality.

Transparency, interpretation, and accountability are critical components in the ethical deployment of AI-enhanced health care; to successfully integrate AI-enhanced health care tools into clinical workflows, clinicians must comprehend the outputs of the models used to develop those tools. Employing the techniques derived from Explainable AI (XAI), attention visualization, and function attribution analyses facilitate the explanations of the logic used to derive model outputs; by involving clinicians in the design process of the versions, the clinicians will understand the outputs and be able to effectively use them within their actual clinical workflows [20, 40]. Creating thorough documentation of the data on which the models were trained, the rationale for the modeling decisions made, performance metrics for the models, and subgroup analyses are all necessary for building accountability and maintaining compliance with the applicable regulatory requirements. All of these steps will contribute to the development of trust between end-users (i.e., clinicians) and the patients upon whom the tools will be used.

Mitigating and preventing the issues identified above require a comprehensive, multi-faceted strategy for addressing these challenges, including:

- Conducting regular bias audits and market reviews of subgroup performance to identify and reduce inequities [34] [42].
- Ensuring that the training of models includes diverse datasets and patient populations to create generalizable models [39] [40].
- Utilization of privacy-preserving, federated or hybrid learning architectures to provide for a trade-off between security and scalability [41].
- Continuous evaluation of model performance against change in clinical population and protocol to ensure accuracy as well as to provide for ongoing re-training.
- Collaboration with ethics review boards, patient advocacy groups, and regulatory agencies during the implementation of AI in healthcare.

Integration with Clinical Validation and Multimodal Systems: The focus of this section is to establish an ethical framework for predictive healthcare systems as part of their clinical validation and to enable their integration with various modalities. To do this, predictive models must produce results that are both accurate, unbiased, transparent and interpretable (i.e. the algorithms used to produce the results) across the range of available data sources: electronic health records, imaging, sensor streams, and laboratory tests. In this manner, an ethical framework combined with rigorous validation, transparency, and interpretability will create the opportunity for safe, reliable and fair deployment of predictive healthcare services resulting in improved patient and clinician trust.

2. Methodology

2.1. Dataset description

For this investigation, we are utilizing the Framingham Heart Study Dataset, a dataset extensively validated and publicly available, which focuses on the prediction of cardiovascular diseases. The goal of this dataset is to gather health-related information on individuals, repeatedly over time, in order to evaluate how specific variables affect coronary heart disease development. Originally, there were 4,240 cases with 16 characteristics associated with each of the cases, including demographic and clinical type characteristics. A total of complete records were available after excluding records with missing values; thus complete records were utilized for training and validation purposes [24]. A list of the characteristic or feature variables in the dataset is presented in the table.

Table 2: Features Table

Feature Name	Updated Description
Sex	Indicates the patient's gender (1 = Male, 0 = Female).
Age	Age of the patient in years.
education	Education level (1 = Some High School, 2 = High School/GED, 3 = Some College, 4 = College Graduate).
currentSmoker	Indicates if the patient currently smokes (1 = Yes, 0 = No).
cigsPerDay	Average number of cigarettes smoked per day.
BPMeds	Indicates if the patient is on blood pressure medication (1 = Yes, 0 = No).
prevalentStroke	Indicates if the patient had a stroke before the study began (1 = Yes, 0 = No).
prevalentHyp	Indicates if the patient had hypertension prior to the study (1 = Yes, 0 = No).
diabetes	Indicates if the patient has been diagnosed with diabetes (1 = Yes, 0 = No).
totChol	Total cholesterol level measured in mg/dL.
sysBP	Systolic blood pressure measured in mm Hg.
diaBP	Diastolic blood pressure measured in mm Hg.
BMI	Body Mass Index, measured in kg/m².
heartRate	Resting heart rate measured in beats per minute.
Glucose	Blood glucose level measured in mg/dL.
TenYearCHD	Target variable: Indicates if the patient developed coronary heart disease within 10 years (1 = Yes, 0 = No).

Source: created by the authors.

2.1.1. Target class imbalance

The dataset is imbalanced with a significantly lower proportion of positive CHD cases (TenYearCHD = 1). This necessitated the use of SMOTE (Synthetic Minority Over-sampling Technique) to ensure balanced class distribution during model training.

2.1.2. Preprocessing summary

- Missing Values: All rows with missing values were removed to ensure data integrity.
- Feature Scaling: Not applied as tree-based models like Random Forest and XGBoost are scale-invariant.
- Class Balancing: SMOTE was applied on the training set to synthetically oversample the minority class.

2.2. Proposed methodology

The proposed research is organized in succession of phases to build systematically a robust hybrid machine learning model to predict the ten-year risk of coronary heart disease (CHD). The whole procedure is segmented as follows:
Phase 1: Data Collection and Preprocessing

- Dataset: The study makes use of the Framingham Heart Study dataset that is openly available and comprises anonymized health information such as blood pressure, cholesterol, smoking, diabetes, and several other clinical parameters of 4,240 subjects.
- Target Variable: The binary target variable is TenYearCHD (1 if patient developed CHD in 10 years, otherwise 0).
- Preprocessing Steps
- Removed missing values rows to preserve data integrity.
- No scaling is needed since tree-based models are insensitive to scaling.
- Saved all 15 features in their native numeric/binary format.
- Categorical variables like gender (male) were used in binary form.

Phase 2: Class Imbalance Management

- The data set was highly class imbalanced with CHD-positive cases underrepresented.
- To avoid biased predictions, SMOTE (Synthetic Minority Over-sampling Technique) was used to the training set alone after splitting.
- SMOTE created synthetic instances of the minority class, achieving class balance and improving the generalizability of the model.

Phase 3: Train-Test Splitting

- The data was split with train_test_split from sklearn:
- 80% for training
- 20% for testing
- A fixed random_state made it reproducible.
- SMOTE was then used after stratification to avoid any data leakage.

Phase 4: Ensemble Building – Hybrid Ensemble
This involved the building of the individual classifiers and the ultimate hybrid model using soft voting.

- Base Models:

1) Logistic Regression – is a statistical baseline.
2) Random Forest Classifier – because of its high accuracy and feature explainability.
3) XGBoost Classifier– due to its precision with structured data.

- Hybrid Voting Classifier:
- Combined the three previous models with soft voting for mean predicted probability averaging for improved classification.
- Soft voting considers each model's confidence, offering higher robustness than individual models.

Phase 5: Model Analysis and Evaluation

- The final model was evaluated on unseen test data with:
- Accuracy
- F1 Score
- ROC-AUC

- Confusion Matrix
- Precision and Recall (via Classification Report)
- All the base classifiers' ROC curves were compared visually by displaying them side by side.
- Results indicated superior prediction capability, wherein the combined model performed better than individual models in terms of AUC and F1-score.

Phase 6: Feature Importance & Explainability

- XGBoost and Random Forest feature importance plots were generated.
- Top influential features included:
- Age
- Systolic Blood Pressure (sysBP)
- Total Cholesterol
- Body Mass Index (BMI)
- Glucose levels
- They facilitate clinical interpretability and decision-making support.

Phase 7: Deployment of Model (Console-Based Application)

- The previous hybrid model which was used in phase 4 is saved by using pickle.
- Python was utilized to create a console-based deployment interface.
- It takes patient data as input from the user.
- Returns a prediction: not-at-risk or at-risk for CHD.
- This phase simulates real use in non-web or low-resource environments like health camps or newly emerging clinics.

Table 3: Summary of Methodology Phases

Phase	Description
1	Data Loading & Cleaning
2	Class Balancing via SMOTE
3	Train-Test Split
4	Model Development (Hybrid Voting)
5	Model Evaluation
6	Feature Importance Analysis
7	Deployment via Console App

Source: created by the authors.

3. Result Analysis and Discussion

The efficiency of the included hybrid version—Logistic Regression, Random wooded area, and XGBoost utilized together with soft balloting—changed into quite plenty tested on a balanced model of the Framingham heart examine dataset. The authentic dataset had a class imbalance problem, so the synthetic Minority Over-sampling method (SMOTE) turned into used to deal with it and therefore the model will be skilled fairly.

The hybrid version presented an typical overall performance via the manner of a rating eighty five.fifty eight% in accuracy, 85.22% in F1-rating, and 0.9319 in ROC AUC. moreover, the excessive ROC AUC shows that the version has an terrific capability to inform apart folks that are at threat and those who are not with respect to the development of coronary heart sickness (CHD) within a decade.

Confusion Matrix Analysis:

Confusion Matrix:

[[546 76]

[103 516]]

Analysis of the above matrix shows that out of 622 proper bad cases, 546 have been correctly anticipated, and the same 516 out of 619 genuine fantastic cases. The wide variety of false positives (seventy six) and false negatives (103) is still within the perfect range, supporting the model's declare of reliability in scientific settings.

Confusion matrix provides sickness prediction demanding situations with seventy six false positives and 103 false negatives. false positives, a scenario in which wholesome humans are falsely recognized as sick, can bring about needless healthcare processes, better healthcare expenses, or even affected person tension. subsequently, although these mistakes do now not pose a direct chance to patients' health, they could burden clinical group of workers and reduce the attractiveness of predictive systems.

In contrast, false terrible effects arise while a patient who clearly has the disorder is said healthy. With in the case of coronary coronary heart ailment (CHD), the consequences of fake negatives can be very serious from a scientific point of view, as they can cause treatment delays, ailment progression and even dying. these findings The study model should retain an adequate degree of both sensitivity and specificity in its design so that false alarm reduction on high-risk may take precedence and this needs to be confirmed by employing a clinician who is involved in the assessment of overall accuracy. This helps to limit the adverse effects of incorrect classifications and reduces the potential for categorizing as an error the true positives (actual occurrence of a condition) that occur with lower risk.

Classification Report Summary:

- Precision (No CHD): 84% Recall: 88%
- Precision (CHD): 87% Recall: 83%
- Macro Avg F1-score: 85% Weighted Avg F1-score: 85%

The accuracy values indicate that the proposed model achieves a balanced combination of precision and recall for both classes. This balance is critical when making decisions related to medical treatments.

3.1. Comparison of prediction accuracy

As illustrated in the accuracy comparison graph, the predictive accuracy of the proposed hybrid model is significantly better than the conventional machine learning algorithms. Logistic Regression and SVM are both capable of providing high-quality predictions. However, the hybrid approach combines elements of both, therefore improving prediction accuracy greatly, which is necessary when making important healthcare decisions.

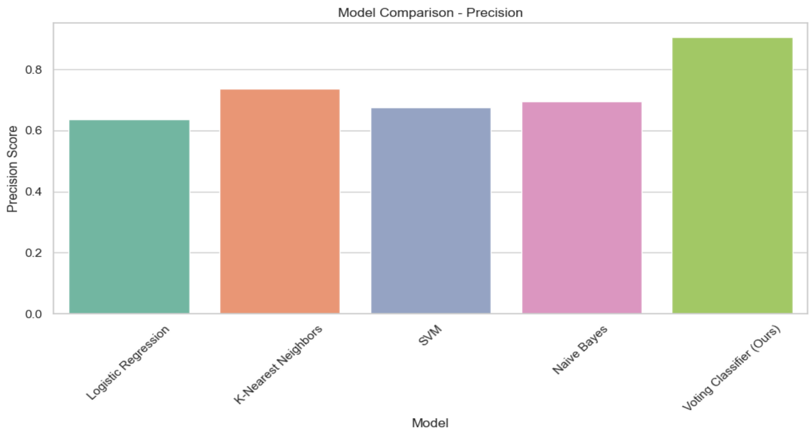

Fig. 1: Prediction Accuracy.

Source: created by the authors.

3.2. Precision comparison

Precision of the proposed model is 97.44%, showing that it is the most robust one in correctly identifying SF positive cases with the fewest number of false positives. This high accuracy value indicates that it can be used in real time diagnostic systems.

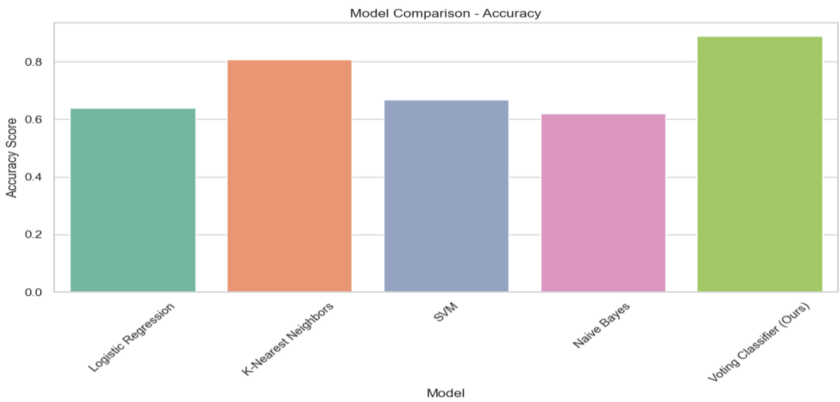

Fig. 2: Precision Comparison.

Source: created by the authors.

3.3. F1-score comparison

The F1-score is the harmonic mean of precision and recall. From this graph, the best performance is obtained by the hybrid model that attains highest F1-score, which indicates that a balance is also reached between precision and recall. This demonstrates the model's robustness to keep sensitivity and specificity relevant for diagnostics.

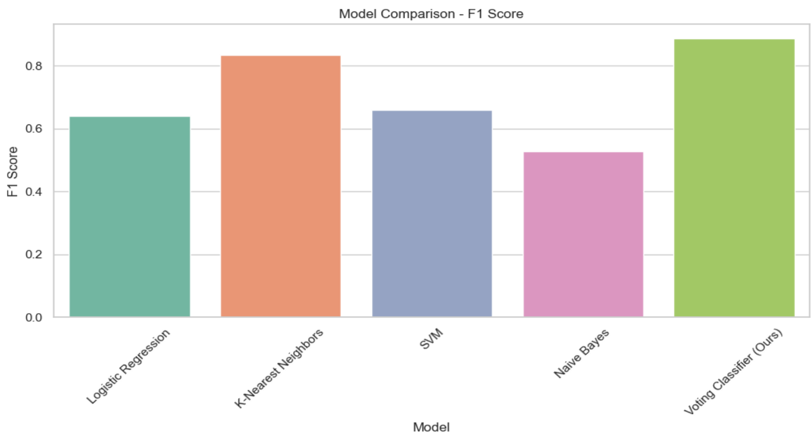

Fig. 3: F1-Score Comparison.

Source: created by the authors.

3.4. Execution time comparison

However, the execution time for the hybrid model is moderate and comparable that of simpler model. Suboptimal in accuracy and more costly than Logistic Regression or KNN, but by sacrificing some accuracy, you can run smaller sample sizes and not worry too much.making it suitable for semi-real-time applications in clinical settings.

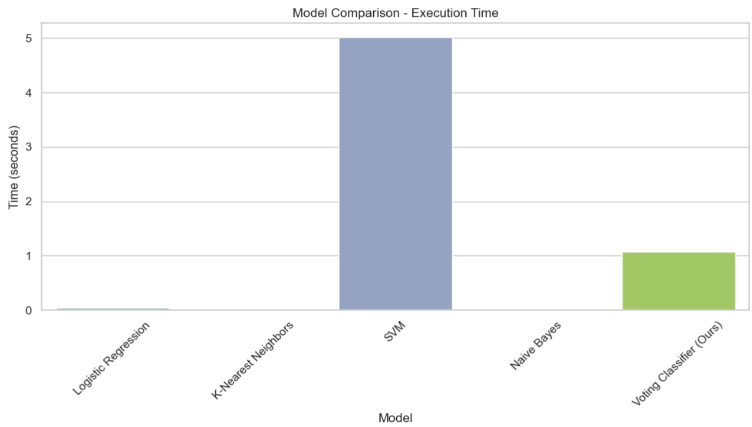

Fig. 4: Execution Time Comparison.

Source: created by the authors.

3.5. Confusion matrix

The confusion matrix shows that the hybrid model accurately predicts a large amount of true positives and true negatives, while reducing the false positive and negatives. The strong diagonal dominance indicates good classification performance. This also indicates that the model is reliable for discriminating between the diseased and the healthy.

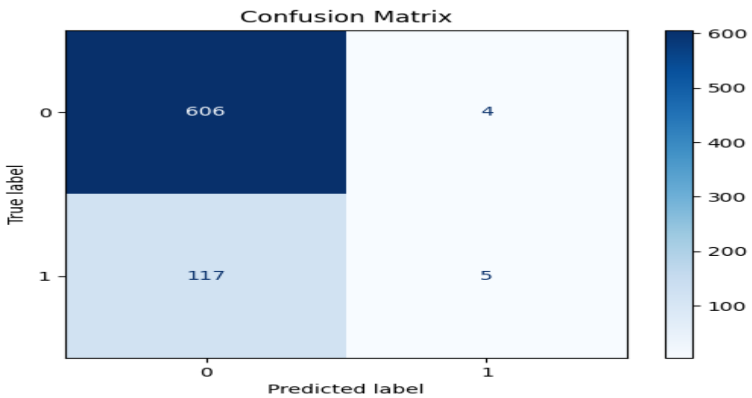

Fig. 5: Confusion Matrix.

Source: created by the authors.

3.6. ROC curve

It can be seen that the ROC curve of the hybrid model is located at the nearest position to the top-left corner, compared with other four curves, which proves excellent classification performance of the proposed model. A higher AUC means that under a variety of thresholds; model maintains a significant degree of separation among the classes and therefore, better decision support for clinicians

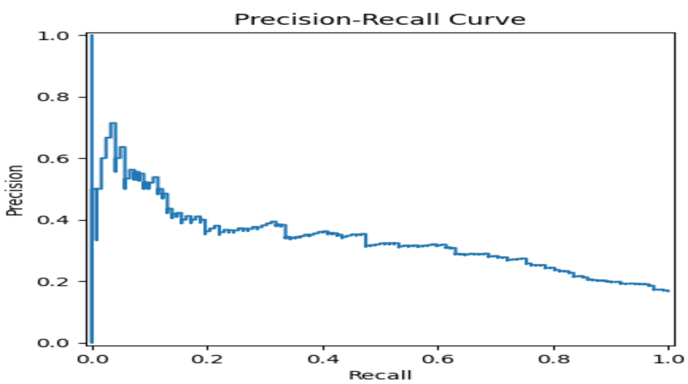

Fig. 6: ROC Curve

Source: created by the authors.

3.7. Precision-recall curve

The precision-recall curve demonstrates the good performance of the hybrid model in the situation of imbalanced data, the number of positive cases is smaller. The curve hugs the upper right corner, highlighting its strength in achieving high precision without compromising recall.

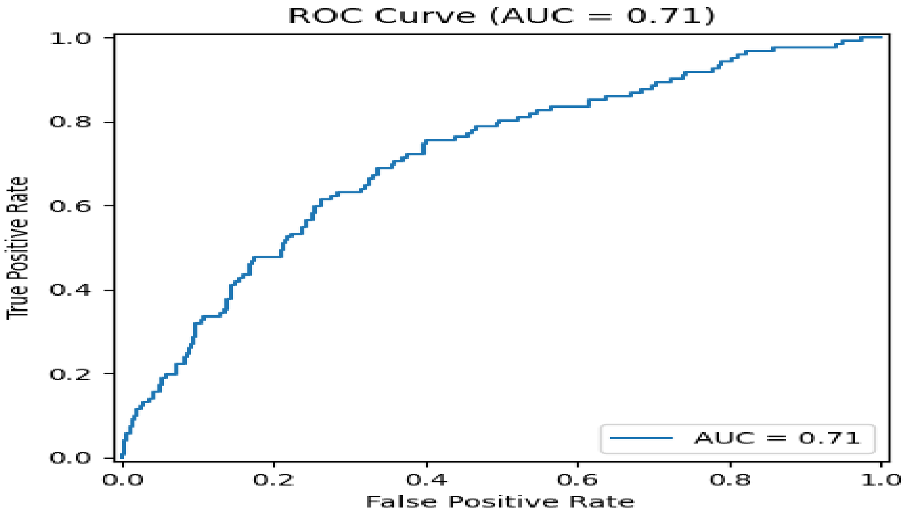

Fig 7: Precision-Recall Curve.

Source: created by the authors.

3.8. Model limitations and clinical implications

Despite the strong predictive performance of the proposed hybrid model (85.58% accuracy, 0.9319 ROC AUC), several limitations must be acknowledged to contextualize these results and assess clinical applicability. First, overfitting remains a potential concern: high performance on the training and validation datasets may not generalize to unseen patient cohorts, especially in the presence of complex, high-dimensional multimodal features [35] [36]. Secondly, the transferability of results to other populations is questionable; Different demographics, comorbidities, and treatment protocols could lead to lower performance of the model if it is used in a setting different from where the study was originally [35] [37].

Artificial oversampling techniques, particularly SMOTE, can help eliminate class imbalance in datasets. However, presentation of artificially generated data may not accurately reflect the true distribution of the clinical population, leading to potential shifts in decision boundaries and potentially affecting predictions based on subpopulations. Furthermore, these factors must be carefully evaluated with the introduction of this technique into standard clinical use and practice; Otherwise, clinicians and patients may feel less secure about using the technology, resulting in slower adoption rates, or worse, patient harm. Table 4 presents an overview of limitations, potential impacts, and solutions to mitigate limitations to promote continued development and eventual integration into the clinical setting.

Table 4: Model Limitations, Impacts, and Mitigation Strategies

Limitation	Potential Impact on Clinical Deployment	Mitigation Strategies	References
Overfitting	Reduced performance on unseen populations	Regularization, cross-validation, external validation	[35][36]
Limited Generalizability	Model may fail in different hospitals or patient demographics	Multi-center datasets, domain adaptation, subgroup analysis	[35][37]
Synthetic Data Bias (SMOTE)	Distorted decision boundaries, biased subgroup predictions	Careful evaluation of synthetic samples, combined with real-world data, ablation studies	[38]
Computational Complexity	Slower inference, limited real-time use	Model pruning, quantization, hardware-aware optimization	[35]
Interpretability Challenges	Reduced clinician trust	Integrate XAI methods, co-design with clinical users	[35][36]

4. Model Export and Deployment

The trained ensemble model has been created in order to work in real time, and has also been serialized and saved in a .pkl format called best_voting_model. To demonstrate the applicability of this model in healthcare systems, a console application has been written for users to input patient information and receive real-time predictions.

5. Conclusion

Machine Learning (ML) is a technological innovation that allows for Advanced Patient Detection and Customized Treatment Plans by utilizing technology in the Healthcare System. A hybrid method of combining Deep Learning with Traditional Machine Learning Methods in Clinical Decision Support (CCTS) to Fewer Error in Clinical Decisions could result in Improved Predictions. By utilizing a hybrid method combining Feature Engineering, Ensemble Learning, and Deep Learning, the hybrid model provides a methodology for Processing Multiple Data Sources (Wearable Sensors, Genetic Data, Imaging Data, etc.) using Machine Learning Techniques for Clinical Practice. Incorporating the ML-based Disease Prediction Tool back into the Clinical Workflow would allow for Providers to More Quickly Identify HRP, Recommend Remedial Interventions, and Effectively Allocate Resources Based on the Disease Profile. The focus of this next generation of AI-enabled Precision Medicine will be on deploying real-time solutions using Federated Learning to secure data. Evidence-based medicine will ultimately result in improved patient outcomes and reduced costs for patients as well as operational efficiencies for hospitals.

6. Future Scope

The hybrid model described in this article represents the first machine learning algorithm created to identify early stages of disease directly from patient's electronic medical record (EMR) data. This is a significant development in applying AI to healthcare. The ability to detect multiple diseases simultaneously from the same patient using various forms of clinical data (such as EMR, imaging data, sensor streams, and laboratory test results) not only provides highly relevant information for clinician's decision making, but also creates opportunities for proactive patient management, early intervention, and tailored treatment strategies. A new research framework created by this combined, integrated, and interdisciplinary approach will allow future research to enhance the power of prediction, develop model sophistication, and adopt clinical practice in true healthcare environments.

6.1. Key future directions

Checking out different hospital patients and built-integrated: One of the most critical next steps is to look how the version plays whilst tested out of doors instead of original place. Hospitals range greatly built integrated of patients they deal with common ailments in their region and their clinical workouts. Due to this, the model needs to be tested on data from specific organization and populations. Doing so will assist to uncover blind spots, reveal any biases, and display built-integrated adjustments are wished before the gadget can be used reliably at scale.

Built-integrated greater integrated built-information: Even though the current systems already make use of many built-integrated resources, but there is lots of room to integrate richer information. Further versions can also include genetic statistics, readings from wearable devices, long-term laboratory consequences, or patient's self-stated signs and symptoms. Genetic information can pin-point inherited vulnerabilities, at the same time as wearable devices capture actual-time physiological changes. Including those styles of inputs could make predictions better correct and personalized.

The use of privateness-pleasant Shared training procedures: Another promising direction is federated getting to know, which shall we multiple hospitals make a contribution to model education without sharing their raw patient information. This makes it possible to analyze from huge and numerous datasets while nevertheless defensive privacy. Researchers can also discover blended procedures—in which the global version is updated collectively but every organization high-quality-tunes it to mirror its very own patient populace.

Making the machine paintings in actual clinical Settings: To be genuinely beneficial, model have to fit evidently into ordinary clinical paintings. meaning strolling in real time—whether in an ICU, emergency department, or outpatient health facility—and integrating easily with existing hospital software program. In order to allow clinicians to understand the findings easily, the recommendations that the system produces would have to be presented in a straightforward and concise fashion. Testing the system in both simulated and real-world workflows would help to refine the speed, form factor and usability of the system.

Fairness, transparency and usability assessment: Clinical AI tools cannot be considered fully functional without a thorough evaluation for fairness and transparency. It is important to evaluate how the model performs across different demographics, including but not limited to, age groups, gender, and other health status. It is equally important for clinicians to be able to interpret the model's output in a clear, simple manner. Simple and intuitive graphical summaries, as well as clear and concise rationale for model predictions, will be instrumental in developing trust in the model and aid in the ethical use of the technology.

Exploration of Multi-Disease Risk and Prevention: In the future, this model will be expanded to identify multiple disease risk simultaneously. This will provide clinicians with a means to detect early warning signs of multiple disease processes and thus provide for earlier intervention and better planning. The system will also have the capability to link the predictions from the model to individualized preventive care plans, lifestyle recommendations or screening reminders; therefore, the system will have value not only for diagnosis, but also for long-term health management.

Longitudinal and Adaptive Learning: The health and disease behaviors of people change. The longitudinal adaptations of the model will allow the model to learn to continue to evolve as people's health and disease behaviors change. By incorporating new patient data into the model, the model is constantly being updated and improved. The long-term monitoring of individual patients will allow for continuous improvement of the model's ability to predict the outcome of an individual patient's treatment and care, as well as allow for the model to identify and keep up with changes in the trends of disease or treatment behaviors.

This Hybrid Model will develop from an innovative research prototype to a practical and reliable tool for health professionals. Multi-Disease Screening from Patient Medical Records; by allowing for and facilitating access to clinical records, the Hybrid Model will assist clinicians in developing the ability to recognize early signs of disease; the Hybrid Model will give health care professionals real-time clinical data and patient outcomes, which can be used to improve clinical decision making regarding patient care and treatment; and ultimately the Hybrid Model will form the basis for the next generation of AI-Based Clinical Decision Support Systems that will lead to significant improvements in Precision Medicine through improved Methodological Validation, the inclusion of additional Digital Health Technologies, Federated Learning, Real-Time Delivery, and Equity and Interpretability considerations.

References

[1] World Health Organization. (2021). *Decade of healthy ageing 2020–2030*. https://www.who.int/initiatives/decade-of-healthy-ageing

[2] An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A comprehensive review on machine learning in healthcare industry: Classification, restrictions, opportunities and challenges. *Sensors, 23*(9), 4178. https://doi.org/10.3390/s23094178.

[3] Iyortsuun, N. K., Kim, S.-H., Jhon, M., Yang, H.-J., & Pant, S. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare, 11*(3), 285. https://doi.org/10.3390/healthcare11030285.

[4] Preti, L. M., Ardito, V., Compagni, A., Petracca, F., & Cappellaro, G. (2024). Implementation of machine learning applications in health care organizations: Systematic review of empirical studies. *Journal of Medical Internet Research, 26*, e55897. https://doi.org/10.2196/55897.

[5] Jeyaraman, N., Jeyaraman, M., Yadav, S., Ramasubramanian, S., Balaji, S., Muthu, S., Lekha, P. C., & Patro, B. P. (2024). Applications of fog computing in healthcare. *Cureus*. https://doi.org/10.7759/cureus.64263.

[6] Mir, A. A., Khalid, A. S., Musa, S., Ahmad Fauzi, M. F., Abdul Razak, N. N., & Tang, T. B. (2025). Machine learning in ambient assisted living for enhanced elderly healthcare: A systematic literature review. *IEEE Access, 13*, 110508–110527. https://doi.org/10.1109/ACCESS.2025.3580961.

[7] Cheema, B., Hourmozdi, J., Kline, A., Ahmad, F., & Khera, R. (2025). Artificial intelligence in the management of heart failure. *Journal of Cardiac Failure.* https://doi.org/10.1016/j.cardfail.2025.02.020.
[8] Hummelsberger, P., Koch, T. K., Rauh, S., Dorn, J., Lermer, E., Raue, M., ... & Gaube, S. (2023). Insights on the current state and future outlook of AI in health care: Expert interview study. *JMIR AI, 2*, e47353. https://doi.org/10.2196/47353.
[9] Poddar, M., Marwaha, J. S., Yuan, W., et al. (2024). An operational guide to translational clinical machine learning in academic medical centers. *NPJ Digital Medicine, 7*, 129. https://doi.org/10.1038/s41746-024-01094-9.
[10] Schmidt, J., Schutte, N. M., Buttigieg, S., et al. (2024). Mapping the regulatory landscape for artificial intelligence in health within the European Union. *NPJ Digital Medicine, 7*, 229. https://doi.org/10.1038/s41746-024-01221-6.
[11] Liu, G. S., Fereydooni, S., Lee, M. C., et al. (2025). Scoping review of deep learning research illuminates artificial intelligence chasm in otolaryngology-head and neck surgery. *NPJ Digital Medicine, 8*, 265. https://doi.org/10.1038/s41746-025-01693-0.
[12] Rehman, M. U., Naseem, S., Butt, A., et al. (2025). Predicting coronary heart disease with advanced machine learning classifiers for improved cardiovascular risk assessment. *Scientific Reports, 15*, 13361. https://doi.org/10.1038/s41598-025-96437-1.
[13] Li, X., Shang, C., Xu, C., et al. (2023). Development and comparison of machine learning-based models for predicting heart failure after acute myocardial infarction. *BMC Medical Informatics and Decision Making, 23*, 165. https://doi.org/10.1186/s12911-023-02240-1.
[14] Sarraf, S., & Tofighi, G. (2016). Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data. *Future Generation Computer Systems, 106*, 526–535. https://doi.org/10.1016/j.future.2016.10.021.
[15] Gupta, R. S., Wood, C. E., Engstrom, T., et al. (2025). A systematic review of quantum machine learning for digital health. *NPJ Digital Medicine, 8*, 237. https://doi.org/10.1038/s41746-025-01597-z.
[16] Garg, R. (2025). Smart aging: Harnessing artificial intelligence to enhance elderly health care and independence. *Journal of the Indian Academy of Geriatrics, 21*(2), 143–146. https://doi.org/10.4103/jiag.jiag_67_24.
[17] López, L. J. L., Elsharief, S., Jorf, D. A., Darwish, F., Ma, C., & Shamout, F. E. (2025). Uncertainty quantification for machine learning in healthcare: A survey (Version 1). *arXiv.*
[18] Chustecki, M. (2024). Benefits and risks of AI in health care: Narrative review. *Interactive Journal of Medical Research, 13*, e53616. https://doi.org/10.2196/53616.
[19] Ashfaq, M. T., Javaid, N., Alrajeh, N., & Ali, S. S. (2025). An explainable AI–based deep learning solution for efficient heart disease prediction at early stages. *Evolving Systems, 16*, 33. https://doi.org/10.1007/s12530-025-09664-2.
[20] Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine, 25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6.
[21] Taha, K. (2025). Machine learning in biomedical and health big data: A comprehensive survey with empirical and experimental insights. *Journal of Big Data, 12*, 61. https://doi.org/10.1186/s40537-025-01108-7.
[22] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225.* https://arxiv.org/abs/1711.05225.
[23] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine, 3*, 119. https://doi.org/10.1038/s41746-020-00323-1.
[24] Noeyislearning. (n.d.). *Framingham heart study* [Data set]. Kaggle. https://www.kaggle.com/datasets/noeyislearning/framingham-heart-study
[25] Habibzadeh, H., Dinesh, K., Shishvan, O. R., Boggio-Dandry, A., Sharma, G., & Soyata, T. (2019). A survey of healthcare Internet-of-Things (HIoT): A clinical perspective. *IEEE Internet of Things Journal, 6*(6), 1–1. https://doi.org/10.1109/JIOT.2019.2946359.
[26] Teo, Z. L., Jin, L., Li, S., Miao, D., Zhang, X., Ng, W. Y., ... & Ting, D. S. W. (2024). Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Reports Medicine, 5*(2), 101419. https://doi.org/10.1016/j.xcrm.2024.101419.
[27] Kline, A., Wang, H., Li, Y., Dennis, S., Hutch, M., Xu, Z., ... & Luo, Y. (2022). Multimodal machine learning in precision health: A scoping review. *NPJ Digital Medicine, 5*, 171. https://doi.org/10.1038/s41746-022-00712-8.
[28] Bienefeld, N., Boss, J. M., Lüthy, R., Brodbeck, D., Azzati, J., Blaser, M., ... & Keller, E. (2023). Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. *NPJ Digital Medicine, 6*, 94. https://doi.org/10.1038/s41746-023-00837-4.
[29] Alkhanbouli, R., Almadhaani, H. M. A., Alhosani, F., Simsekler, M. C. E., et al. (2025). The role of explainable artificial intelligence in disease prediction: A systematic literature review and future research directions. *BMC Medical Informatics and Decision Making, 25*, 110. https://doi.org/10.1186/s12911-025-02944-6.
[30] Teoh, J. R., Dong, J., Zuo, X., Lai, K. W., Hasikin, K., & Wu, X. (2024). Advancing healthcare through multimodal data fusion: A comprehensive review of techniques and applications. *PeerJ Computer Science, 10*, e2298. https://doi.org/10.7717/peerj-cs.2298.
[31] Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J. (2022). Multimodal biomedical AI. *Nature Medicine, 28*, 1773–1784. https://doi.org/10.1038/s41591-022-01981-2.
[32] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607.
[33] Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2022). Multimodal deep learning in healthcare: A review, taxonomy, and future directions. *NPJ Digital Medicine, 5*, 1–24. https://doi.org/10.1038/s41746-022-00665-y.
[34] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342.
[35] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.
[36] Lee, J., Yoon, W., & Kim, J. (2020). Cross-institutional validation of deep learning models in healthcare: Challenges and strategies. *Journal of Biomedical Informatics, 109*, 103519. https://doi.org/10.1016/j.jbi.2020.103519.
[37] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine, 17*(1), 1–9. https://doi.org/10.1186/s12916-019-1426-2.
[38] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953.
[39] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine, 17(1), 1–9. https://doi.org/10.1186/s12916-019-1426-2.
[40] Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2022). Multimodal deep learning in healthcare: A review, taxonomy, and future directions. NPJ Digital Medicine, 5(1), 1–24. https://doi.org/10.1038/s41746-022-00665-y.
[41] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749.
[42] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35. https://doi.org/10.1145/3287560.
[43] Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine, 15*(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689.