

IJISRT25NOV1164

by A A

Submission date: 25-Nov-2025 07:51PM (UTC+0200)

Submission ID: 2827511209

File name: 1763904852.docx (363.94K)

Word count: 23407

Character count: 147077

Leveraging Machine Learning for Cardiovascular Disease Prediction: A Review of Current Approaches, Challenges, and Opportunities

Abstract— This review explores the application of machine learning (ML) models in cardiovascular disease (CVD) prediction, emphasizing their potential to revolutionize risk assessment and healthcare delivery. By examining traditional models such as Logistic Regression, Support Vector Machines, and Random Forest, alongside advanced approaches like XGBoost, the study highlights their strengths, limitations, and performance in various healthcare contexts. The review underscores the growing role of hybrid and explainable ML architectures, as well as the integration of deep learning techniques, in enhancing accuracy, scalability, and clinical trust. Despite notable advancements, challenges such as data quality, imbalance, and ethical considerations in underrepresented regions persist, underscoring the need for collaborative efforts among stakeholders to ensure equitable and efficient implementation. By addressing these gaps and **raging robust models like XGBoost, ML has the potential to significantly reduce the global burden of CVDs and drive transformative change in precision medicine and patient-centred care.**

Keywords—, Data Science, Big Data, Machine Learning, Artificial Intelligence, Computer Science, Cardiovascular Disease Prediction, Machine Learning Models, Machine Learning Architectures

I. INTRODUCTION

A. Background

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide, accounting for approximately 17.9 million deaths annually, which represents nearly one-third of all global deaths [1]. Notably, over 75% of these fatalities occur in low- and middle-income countries, underscoring a significant health disparity [2].

The global burden of CVDs has been escalating over the past decades. In 2000, around 14 million individuals succumbed to cardiovascular conditions; by 2019, this number had risen to nearly 18 million[3]. This increase is attributed to factors such as population growth, aging demographics, and the prevalence of risk factors like hypertension, obesity, and diabetes. Countries with the highest number of CVD-related deaths include China, India, Russia, the United States, and Indonesia [4].

Examining the continental landscape, Eastern Europe exhibits the highest age-standardized CVD mortality rates,

reaching 432.3 per 100,000 individuals as of 2022[5]. In contrast, regions like the High-income Asia Pacific report significantly lower rates, approximately 73.6 per 100,000.

These disparities highlight the varying impact of CVDs across continents, influenced by factors such as healthcare infrastructure, socioeconomic conditions, and lifestyle behaviours [6].

In Africa, the burden of CVDs is substantial and growing. The continent is home to over 1 billion people and contributes significantly to the global CVD burden. In 2013, an estimated 1 million deaths were attributed to **57 cardiovascular diseases in Africa [7]**. Hypertension is identified as the primary driver of CVDs in the region, with the African continent having the highest prevalence of hypertension globally[8]. Factors such as urbanization, aging populations, social stress, and limited access to healthcare services exacerbate this issue.

Focusing on Uganda, CVDs are a leading health concern. In 2019, the country reported 28,149 deaths due to cardiovascular diseases [9]. This places Uganda in the bottom 40% of countries concerning age-standardized mortality from CVDs. The nation faces challenges such as high rates of air pollution and elevated blood pressure among females compared to global averages[10]. Despite these challenges, Uganda has implemented seven out of eight key CVD-related policies and has a national action plan for cardiovascular diseases in place. However, the absence of an operational unit within the Ministry of Health dedicated to non-communicable diseases indicates areas needing further development [11].

B. Objective of the Review

61 objective of this review is to evaluate the effectiveness of machine learning models in predicting cardiovascular diseases

C. Machine Learning Overview

Machine Learning (ML), a subset of artificial intelligence (AI), enables computers to learn from data and make

predictions without explicit programming. By recognizing patterns, ML algorithms [32] excel in classification and decision-making across various fields, including natural language processing, computer vision, and robotics. [12] ML is broadly categorized into supervised learning, unsupervised learning, and reinforcement learning, each serving distinct purposes. Its success depends on [53] quality and quantity of training data. ML is transforming industries such as healthcare, finance, transportation, and e-commerce, automating processes and enhancing decision-making [12].

[34] However, ML has challenges. Issues such as ensuring the quality of data, making models interpretable, and addressing ethical concerns remain pressing [13]. These challenges underscore the need for adaptation, especially in resource-limited areas like Uganda. Despite obstacles, ML can drive innovation by enabling efficient data analysis, improving predictions, and supporting informed decision-making, delivering transformative benefits to communities and industries.

II. MACHINE LEARNING ARCHITECTURES OF CVD PREDICTION

Machine learning (ML) architectures process data to enable automation and data-driven insights. These range from Traditional ML models like decision trees and linear regression, valued for efficiency in low-data environments [14], to Deep Learning architectures, which use neural networks for high-dimensional tasks but require significant computational power [15]. Hybrid ML architectures combine traditional and deep learning models, enhancing scalability for complex applications [16]. Explainable AI (XAI) architectures improve transparency, crucial in sensitive fields like healthcare and finance [17]. Multi-modal Learning architectures integrate diverse data types, such as text and images, to enhance accuracy in tasks like NLP and autonomous systems [17]. As ML advances, balancing efficiency, interpretability, and computational demands remains a key challenge.

A. TRADITIONAL MACHINE LEARNING ARCHITECTURES

[49]

Traditional machine learning (ML) Architectures have played a significant role in predicting cardiovascular diseases (CVD) by utilizing structured medical datasets. This section delves into the commonly used ML approaches, their applications in CVD prediction, notable studies that have implemented these methods, and the strengths and limitations of each architecture. [18]

Machine learning techniques such as Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, Naïve Bayes, k-Nearest Neighbors, and Gradient Boosting Machines have been widely applied to this field, yielding varying degrees of success depending on the context. By reviewing recent advancements, this discussion sheds light on how these traditional ML architectures have

been used, their practical benefits, and the challenges they face in accurately predicting and managing cardiovascular disease. [18]

Logistic Regression (LR)

Logistic Regression (LR) is a widely used machine learning model in healthcare, particularly for binary classification tasks like determining whether a patient is at risk for cardiovascular disease (CVD). A notable example is a study by [19], which employed LR to estimate 10-year CVD risk based on routine physical examination data. [19] The model is favoured in clinical settings for its simplicity and high interpretability, allowing healthcare professionals to clearly understand how individual risk factors contribute to predictions. However, its reliance on linear relationships between predictors and outcomes can limit its ability to capture more complex interactions. Additionally, it may struggle when applied to high-dimensional datasets or scenarios involving non-linear patterns.

Decision Trees (DT)

Decision Trees (DTs) are widely appreciated for their intuitive visualization and straightforward decision-making process, making them especially useful in clinical settings. For example, a study in South Korea successfully utilized DTs to classify cardiovascular disease (CVD) risk factors, showcasing their ability to pinpoint critical predictors [20]. These models are not only easy to interpret but also capable of capturing non-linear relationships between variables and outcomes. However, they come with challenges, such as a tendency to overfit, particularly with smaller datasets, and sensitivity to minor data changes, which can drastically alter the tree structure. [20]

Random Forests (RF)

Random Forest (RF) is a powerful ensemble learning method that improves disease prediction by analysing complex medical data. It enhances classification accuracy by combining multiple decision trees, reducing overfitting, and ensuring reliable outputs. RF has been highly effective in diagnosing conditions like cardiovascular diseases and diabetes, achieving up to **91% accuracy** in genomic analysis and healthcare diagnostics [21] [22]. Additionally, it highlights key disease predictors, aiding medical research and decision-making [23]. However, RF demands high computational power, making it resource-intensive for large datasets [22]. Its ensemble structure also reduces interpretability, posing challenges in clinical applications where transparency is essential [24]. While simpler models may sometimes be preferable for clarity, RF remains a top choice for disease prediction due to its accuracy and robustness [23].

30

Support Vector Machines (SVM)

Support Vector Machines (SVMs) are widely used in disease prediction due to their ability to classify medical data by identifying the optimal hyperplane that separates different health conditions. Their strength lies in handling high-dimensional datasets, making them particularly useful for analysing genetic and metabolic markers[25]. Additionally, SVMs reduce overfitting risks by maximizing the margin between classes, ensuring stable predictions even with limited data[26]. Their adaptability [27] through kernel functions also allows them to model complex, non-linear relationships in disease diagnostics [25]. However, SVMs require careful hyperparameter tuning, which can be time-consuming and affects model performance [27]. Their high computational costs limit scalability for large medical datasets, making them less suitable for real-time applications [28].

Additionally, their black-box nature makes interpretation challenging, reducing transparency in clinical decision-making[27]. While SVMs remain a strong choice for disease prediction, neural networks may offer better interpretability and scalability for specific healthcare needs.

17

Naïve Bayes (NB)

Naïve Bayes is a probabilistic machine learning algorithm that classifies data based on conditional independence, making it effective for disease prediction. It has been successfully applied to heart disease, diabetes, chronic kidney disease, and cancer, achieving 85% accuracy in heart disease prediction and 95% in intelligent health risk assessment [29]. [28] Its efficiency, ease of implementation, and ability to handle large datasets make it a valuable tool in medical diagnostics [31]. Additionally, it can sometimes outperform more complex models, proving its reliability in real-world healthcare applications [32]. However, its assumption that features are independent can reduce accuracy when dealing with correlated medical variables [32]. It also faces the zero-probability problem, which requires techniques like Laplace smoothing to improve prediction reliability [28]. Despite these limitations, Naïve Bayes remains a fast, scalable, and effective algorithm for disease classification.

K-Nearest Neighbours(K-NN)

K-NN is a non-parametric algorithm that classifies data based on proximity to labelled points, making it useful in disease prediction. In heart disease detection, it evaluates factors like age, gender, and medical history, outperforming traditional methods[30]. It also aids breast cancer diagnosis by classifying tumors as benign or malignant based on attributes like texture and perimeter. k-Fold Cross Validation (k-FCV) helps optimize k for better accuracy [31]. Beyond diagnosis, K-NN supports personalized health supplement recommendations

by matching disease profiles with functional foods [33]

K-NN's versatility enables applications across medical fields [32]. Its non-parametric nature suits real-world datasets but makes it computationally expensive, requiring distance calculations for each input. Sensitivity to noise and irrelevant features can affect predictions, and performance depends on choosing the right k-value and distance metric [34]. Despite these challenges, careful preprocessing and parameter tuning enhance its effectiveness in disease prediction.

Gradient Boosting Machines (GBM)

XGBoost is a powerful machine learning algorithm that enhances prediction accuracy by combining multiple weak models iteratively. In heart disease prediction, XGBoost outperformed other models by analysing key indicators such as fasting blood sugar and cholesterol[35]. Similarly, in Parkinson's disease, it provided superior accuracy over traditional diagnostic methods, effectively tracking disease progression[36]. The integration of boosting techniques like AdaBoost and Light GBM has further improved performance, with some models achieving up to 99.75% accuracy [37]. XGBoost is valued for its high accuracy, flexibility, and ability to capture complex patterns in medical data [37]. It can process diverse datasets, including clinical records and genetic data, making it versatile for different diseases [38]. However, its performance depends on data quality, limiting effectiveness in cases of missing or biased data [38]. Additionally, its complexity makes interpretation difficult, posing challenges for clinical integration [38]. XGBoost is also computationally intensive, requiring significant resources [36]. Despite these challenges, its predictive power makes it a valuable tool in disease forecasting, provided efforts are made to improve interpretability and data handling.

B. DEEP LEARNING ARCHITECTURES

Deep learning (DL) architectures have emerged as powerful tools for the prediction and diagnosis of cardiovascular diseases (CVDs), leveraging their ability to process complex, high-dimensional, and multi-modal data. This section reviews the use of various deep learning models in CVD prediction, highlighting their applications, strengths, weaknesses, and contributions in the health sector at large.

11

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are deep learning models designed for processing complex image data, making them highly effective in disease prediction. They have been applied in diagnosing conditions like Parkinson's and malaria by analysing MRI scans and cell images, enabling early and accurate detection [39][40]. CNNs also support home-based disease detection, allowing individuals to monitor their health remotely [41]. Their ability to extract spatial features makes them ideal for medical imaging, with models like VGG16 and ResNet50 achieving accuracy rates of up to 97% in Parkinson's diagnosis[42][39].

Advanced techniques such as dropout and batch normalization further improve precision and reliability [40]. Despite their effectiveness, CNNs face challenges in healthcare. They require large datasets and high computational power, limiting accessibility [42]. Privacy concerns and data quality issues hinder adoption, while mathematical validation is crucial for reliability [43]. Overcoming these challenges with ethical use and data security will enhance their impact in disease prediction.

1

Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are deep learning models designed to process sequential data, making them ideal for analyzing time-series information in healthcare. They are used to predict disease progression by analyzing patient history and genetic profiles, improving early detection and risk assessments [44],[45]. Their ability to capture temporal dependencies allows for more precise prognostication compared to traditional methods, enhancing diagnostic accuracy and accessibility [44]. RNNs also adapt well across different healthcare applications, making them valuable for medical data analysis [45].

Despite their benefits, RNNs face challenges. They require significant computational resources for training and deployment [45] and are sensitive to data quality and privacy issues, which can limit broader adoption [43]. Traditional RNNs also struggle with the vanishing gradient problem, affecting their ability to learn long-term dependencies[46]. Addressing these issues through hybrid models, such as combining RNNs with other neural networks, could improve their efficiency and predictive performance in healthcare.

Deep Belief Networks (DBNs) are deep learning models designed for handling high-dimensional data, making them effective in disease prediction. They have been used to optimize feature extraction and classification in healthcare, improving accuracy and efficiency. For instance, an Optimized DBN integrated with CNN-LSTM and UFDSOM achieved 98.23% accuracy in disease classification [47] DBNs have also been applied in Alzheimer's detection using DNA methylation and gene expression data, effectively addressing the High Dimension Low Sample Size issue and outperforming traditional models [48][49]

DBNs offer high accuracy, robust feature extraction, and effective integration with other models like CNNs and LSTMs [47]. However, their training process is complex due to multiple learning variables, often requiring optimization techniques like the Ant Lion Optimization algorithm [47]. They are also sensitive to noise in medical data, necessitating advanced learning algorithms for reliability[50]. Overcoming these challenges through improved training methods and hybrid models could further enhance DBNs' role in disease prediction.

Autoencoders

Autoencoders are unsupervised learning models that extract efficient data representations, making them valuable in medical data analysis and disease prediction. They have been used to enhance feature extraction, improve genetic risk assessment, and predict disease associations. The Autoencoder-based Broad Learning model combines denoising autoencoders with incremental learning, achieving 98.50% accuracy in dynamic healthcare settings [51]. Genetic autoencoders help predict disease-associated single nucleotide polymorphism, improving risk scores for conditions like multiple sclerosis [52]. The End-to-End Supervised Autoencoder model surpasses traditional methods in predicting co-existing diseases using diagnostic codes[53] , while the SARMDA model integrates adversarial regularization with graph neural networks to predict microbe-disease associations effectively [54] Autoencoders excel in feature extraction, noise reduction, and high prediction accuracy, particularly in complex datasets[51],[52]. However, they require significant computational resources, especially in genetic analysis [52], and may struggle to capture global patterns without advanced regularization[54]. Addressing these challenges through optimized learning algorithms will further enhance their role in disease prediction.

Deep Belief Learning Architectures

Transformers

Transformers are deep learning models known for processing complex data and identifying patterns, making them highly effective in disease prediction. They have been adapted for diagnosis, prognosis, and clinical outcome forecasting. DiagnoAI, for instance, utilizes a pre-trained transformer to predict diseases from symptom descriptions, excelling in multilingual settings[55]. TransformEHR predicts disease outcomes from electronic health records, aiding in early detection of conditions like pancreatic cancer [56]. OncoBERT leverages unstructured EHR notes for cancer outcome prediction, improving both accuracy and interpretability[57]. DeepProg focuses on disease progression analysis using images and auxiliary data[58], while transformers have also been explored for immunotherapy prognosis, enhancing early disease detection[59]. Transformers offer high accuracy, adaptability across different diseases, and improved interpretability, as seen with OncoBERT's transparency in decision-making [57] However, they require large datasets, which can be a limitation for rare diseases, and their training demands significant computational resources, restricting accessibility in lower-resource settings[58]. Overcoming these challenges could further expand their role in clinical applications.

Hybrid Deep Learning Architectures

Hybrid deep learning architectures combine multiple neural network models to enhance disease prediction accuracy and efficiency. By leveraging the strengths of different models, these architectures tackle complex medical data challenges. For instance, a hybrid model integrating genetic algorithms, stacked autoencoders, and Softmax classifiers outperformed traditional methods in early diabetes detection, leading to a web-based diagnosis tool [60]. Similarly, a CNN-LSTM model achieved high accuracy in chronic disease prediction, with 98% for cancer and 99% for kidney disease[61]. Another approach used a Coati Optimization Algorithm within an ANN framework, improving cardiovascular disease prediction [62], while a Bi-GRU-LSTM model effectively handled small, complex datasets for diabetes-related heart disease monitoring [63]

Hybrid models excel in accuracy, versatility, and advanced feature extraction, making them highly adaptable across different diseases [60][61]. However, their complexity makes training and optimization challenging[64], and they require large, high-quality datasets, which may not always be available [63]. Simplifying these models and improving efficiency could enhance their accessibility for wider healthcare applications.

C. HYBRID ARCHITECTURES

Hybrid Machine Learning architectures integrate multiple computational models to enhance performance, scalability, and robustness. These architectures address challenges like adversarial attacks, distributed processing, and model efficiency by leveraging specialized layers and hybrid topologies [65]. They improve neural network resilience in tasks like image classification[66], enhance emotion recognition in speech [67], and optimize distributed learning by reducing computational costs in wireless networks [68] as seen below.

Convolutional Neural Networks (CNN) with Recurrent Neural Networks (RNNs)

Hybrid CNN-RNN architectures have proven effective in disease prediction by combining CNNs for spatial feature extraction with RNNs for capturing temporal patterns, enhancing diagnostic accuracy. These models have been applied across various medical domains, including anaemia detection, where a CNN-RNN model achieved 90.27% accuracy, outperforming standalone CNNs and RNNs[69]. Similarly, a CNN-LSTM model demonstrated superior accuracy in predicting diseases from Electronic Health Records compared to SVMs and standalone deep learning models [64] In chronic disease detection, a CNN-LSTM hybrid achieved 98%-99% accuracy for kidney disease[61] while a CNN-RNN model utilizing transfer learning excelled in COVID-19 detection via X-ray analysis [70]. These models offer enhanced feature extraction[70], improved accuracy across multiple datasets[69], and versatility in handling different medical data types[64]. However, they pose challenges such as increased computational complexity, requiring more resources and longer training times [45], and dependence on large, high-quality datasets [70]. Future research could refine these models by integrating attention mechanisms to improve efficiency while maintaining high predictive performance.

Support Vector Machines (SVM) with Neural Networks

Hybrid SVM-NN models enhance disease prediction by leveraging both approaches. A hybrid ANN-SVM model improved diabetes classification to 89.7% accuracy using non-linear kernel SVMs [71]. In medical imaging, SVMs acted as primary classifiers while ANNs refined results, boosting accuracy [72]. The Neural Support Vector Machine outperformed standard SVMs and multi-layer perceptrons in regression and dimensionality reduction tasks[73]. These models offer high accuracy[71] effective dimensionality, and adaptability via hyperparameter tuning. However, they demand high computational resources and require extensive hyperparameter optimization[74]. Despite challenges, SVM-NN hybrids advance disease prediction if computational constraints are managed.

Autoencoders with Decision Trees

Autoencoder-based hybrid models enhance disease prediction by combining feature extraction with

50 interpretability. Autoencoder Trees, introduced by Ozan Irsoy and Ethem Alpaydin, use **decision trees** for encoding and decoding, leveraging soft decision trees with multivariate splits for hierarchical data representation [75]. They achieve strong reconstruction accuracy on datasets like handwritten digits and news data [75]. Encoder Forest, proposed by Ji Feng and Zhi-Hua Zhou, integrated tree ensembles with autoencoders, offering lower reconstruction error and faster training than deep neural networks[76]. Its robustness makes it applicable to disease prediction[76]. Hybrid frameworks, such as those explored by P. Epsoba et al., apply these models to medical prognosis, proving their utility in healthcare[77]. While these architectures improve interpretability and efficiency, they risk overfitting and performance variability based on dataset characteristics. Nonetheless, they provide a promising direction for disease prediction and complex data analysis.

CNNs with Transformers

16 hybrid CNN-Transformer models enhance medical image analysis by combining CNNs' local feature extraction with Transformers' global context modelling, improving disease **18** prediction accuracy. Li Wang et al. developed a CNN-Transformer model for N-staging and survival prediction in non-small cell lung cancer (NSCLC), outperforming standalone CNNs and Transformers[78]. ReFormer, introduced by [79], employs frequency-based bridging to refine feature fusion in medical image segmentation[79]. These models improve predictive accuracy[79] ,optimize feature fusion[79] , and adapt across imaging tasks[80]. However, their complexity and high computational costs pose challenges [81]. Feature misalignment can also hinder effective fusion [79]. Despite these limitations, refining integration strategies can enhance efficiency, making CNN-Transformer hybrids vital for medical diagnostics[81].

D. EXPLAINABLE ARTIFICIAL INTELLIGENCE ARCHITECTURES

Explainable AI (XAI) architectures enhance transparency by making AI decisions interpretable, using techniques like LIME [82]. They improve trust in applications like pathology [83] and cybersecurity[84] but face challenges balancing accuracy and interpretability [85] as seen below.

Explainable Boosted Trees

Explainable Boosted Trees architectures balance predictive accuracy and interpretability, crucial for disease prediction in healthcare. By integrating

52 techniques like LIME [86], EBTs provide insights into decision-making processes, improving trust and clinical decision-making in areas like chronic kidney disease and ocular disease detection [87]. Their strengths include enhanced interpretability[86] ,supporting better communication in healthcare, and high accuracy [87]. However, challenges exist in implementation complexity [87] and limited generalizability across conditions [88].

35 Decision Trees with Rule Based Explanations

Decision trees with rule-based explanations are essential in explainable AI (XAI), particularly in disease prediction, by providing transparent, interpretable decision-making processes. For example, a decision tree model for chronic kidney disease diagnosis uses Gini Importance and SHAP values for interpretation [89]. The SRules method enhances the interpretability of complex models by creating concise, understandable rules[90], and integrated systems like decision trees with random forests use SHAP and LIME for medical predictions, increasing trust [91]. These models improve transparency and decision-making [91] and offer **39** finalization for better predictions[92]. However, there are trade-offs between interpretability and accuracy, and creating surrogate models can be resource-intensive[90].

Explainable Neural Networks

Explainable Neural Networks (XNNs) aim to improve transparency in AI-driven medical diagnostics by integrating explainability techniques into neural network architectures. These models maintain high diagnostic accuracy while providing interpretable outputs, crucial for clinical adoption. For instance, Thakur's research enhances deep learning models with XAI techniques for medical diagnosis [93], and[94] combine large language models with explainability algorithms for better disease pred. **58** is [94] Jere's study uses ensemble learning and **SHAP** for heart disease prediction, achieving high performance [95], while Mathew et al. integrate SHAP and LIME in decision tree models for medicine recommendations [91]. XNNs offer high transparency [95] and robust diagnostic performance [93], empowering clinicians with understandable explanations [91]. However, these models can be complex to implement and may risk overfitting with limited data, along with scalability challenges [96]. Despite these, XNNs provide crucial insights into medical AI systems, promoting trust and understanding.

Bayesian Networks

Bayesian Networks (BNs) are a key architecture in explainable AI (XAI), particularly for disease

prediction. BNs utilize probabilistic graphical models to capture dependencies and causal relationships, making them ideal for medical diagnostics, where uncertainty is common. Researchers like Jose A. Quesada and colleagues have extensively reviewed BNs' use in medical diagnostics over the past 40 years[97]. They are widely used in disease diagnosis and prognosis, with metrics like sensitivity and specificity to assess performance. BNs' strengths include clear explainability, integration of expert knowledge, and effective uncertainty handling, crucial for medical applications[98]. However, they face challenges in managing complexity as the number of variables grows and their unstructured application limits effectiveness [98]. Addressing these challenges, researchers aim to create structured frameworks for better interpretability and communication in BNs [98]

E. MULTI-MODAL LEARNING ARCHITECTURES

Multimodal learning architectures combine data from different modalities (e.g., text, images, audio) to improve predictions. These models capture both inter- and intra-modality relationships, enhancing decision-making [99]. Integration can occur via deep or early fusion [100], but challenges include data heterogeneity and complexity [101]

Deep Neural Networks in Multimodal Learning

DNNs in multimodal learning have revolutionized disease prediction by integrating various data types like omics, imaging, and structured data to improve model accuracy. For example, Stochastic Training with Regularization and Gradient Descent by Ohnuki [102] predicts drug-disease relationships using multi-omics data [102], while [51] ABL model enhances feature extraction in medical data with a Denoising Auto Encoder [51]. Other models like IDC-GN and MF-DL use advanced graph-based techniques and multi-omics data to enhance prediction accuracy [103][104]. These models are strong in integrating diverse data, improving accuracy, and handling noisy data[51]. However, challenges include high computational demands, model complexity, and reliance on high-quality multimodal data. Future work should focus on optimizing these models to balance performance and complexity in diverse healthcare applications.

Graph Neural Networks (GNN)

Graph Neural Networks (GNNs) play a crucial role in multimodal learning for disease prediction by modelling complex relationships across various data types, such as

clinical, demographic, and imaging data [105],[106]. Models like IDC-GN and STRGNN use multi-omics data and advanced graph techniques to enhance prediction accuracy [103],[102]. GNNs are widely used in applications like heart disease prediction and veterinary epidemiology, where they integrate data from diverse sources to improve predictive performance[106],[107]. Their strengths include the ability to capture intricate relationships and scale with large datasets, but they face challenges such as data noise and high computational demands [105],[103]. Future research may focus on refining GNNs for better performance and interpretability.

III. **MACHINE LEARNING MODELS**

Machine learning (ML) models are computational systems that analyse data to identify patterns and make predictions without explicit programming. In cardiovascular disease (CVD) prediction, ML enhances early detection, risk assessment, and personalized treatment by uncovering hidden patterns in complex datasets. Key models include Random Forest, which reduces overfitting for reliable predictions, and Support Vector Machines (SVM), known for handling non-linear data with high accuracy [108], [109]. Logistic Regression, which offers interpretability, K-Nearest Neighbors (KNN) that provides strong classification, and XGBoost excels in processing large datasets [110],[111]. Beyond prediction, ML aids biomarker discovery and improves clinical workflows, though challenges like data quality, interpretability, and ethical concerns must be addressed for effective healthcare integration[112], [113]

i. Support Vector Machine (SVM):

Support Vector Machines are supervised learning models used for classification and regression, excelling in high-dimensional data analysis through kernel functions[114], [115]. In cardiovascular disease (CVD) prediction, SVMs have shown high accuracy across studies. In India, they outperformed models like XGBoost[116], while in China, SVMs with radial basis functions excelled in coronary heart disease classification [117]. Bangladesh researchers found polynomial SVMs more accurate than linear ones [118]. Comparative studies consistently highlight SVM's strong predictive power [119].

Despite their accuracy and robustness, SVMs face challenges such as computational intensity, overfitting risks, and limited interpretability [120],[119]. Optimizing hyperparameters and improving model explainability could enhance their clinical applicability.

ii. XGBoost (eXtreme-Gradient Boost Algorithm):

XGBoost is an ensemble learning algorithm based on decision trees, designed for speed and efficiency in handling large datasets[121], [122]. Widely used in cardiovascular disease (CVD) prediction, it has demonstrated high accuracy

in identifying heart attack and stroke risks. In the UK, researchers developed a mixed-effects XGBoost model to predict 30-day mortality post-cardiac surgery, improving predictions by accounting for hospital-level variations [121]. Other studies found XGBoost outperformed models like Random Forest and Logistic Regression in stroke prediction, highlighting its precision in classifying at-risk individuals[123] .

XGBoost excels in processing complex medical data with superior predictive accuracy, making it more effective than models like KNN and Naïve Bayes[122] . However, it struggles with correlated data without modifications and requires careful hyperparameter tuning, which can be computationally demanding [121]. Optimizing its implementation and integrating clinical insights can enhance its practical use in healthcare.

60

iii. **K-nearest Neighbours (KNN) Classifier:**

K-Nearest Neighbors (KNN) is a supervised learning algorithm used in cardiovascular disease (CVD) prediction to classify individuals based on their risk of heart-related conditions. It works by identifying the 'k' closest data points to an input and predicting outcomes based on the majority class. In a study using Kaggle data, KNN was compared with Logistic Regression and Random Forest, where it performed well but was outmatched by Random Forest in accuracy[124] . In Indonesia, researchers applied KNN with Principal Component Analysis (PCA) for feature selection, achieving 90.16% accuracy in heart attack risk classification, demonstrating its clinical potential [125]. Another study highlighted KNN's role in reducing misdiagnosis and improving patient classification [126].

33

KNN is valued for its simplicity and effectiveness in high-dimensional data when paired with proper feature selection [124],[125]. However, it is computationally intensive for large datasets and sensitive to irrelevant features, which can impact performance if not properly filtered [124],[125]. While KNN is useful in CVD prediction, models like Random Forest and SVM may offer better accuracy depending on the dataset and clinical application.

iv. **Random Forest Classifier:**

Random Forest Classifier is an ensemble machine learning model used for cardiovascular disease (CVD) prediction that analyses multiple risk factors. It constructs numerous decision trees and combines their outputs to enhance accuracy and reduce overfitting. Studies in India and other countries have utilized this model to integrate clinical, demographic, and genetic data for improved CVD risk assessment[127]. The goal is to

enhance sensitivity and specificity, enabling early diagnosis and personalized treatment[128]

1

Random Forest is known for its high accuracy, outperforming models like Logistic Regression and SVM[127]. It also provides insights into feature importance, helping identify key risk factors [129]. However, its complexity makes interpretation difficult, and its performance depends on data quality [128], [129]. Despite these challenges, optimizing the model and integrating additional predictive factors could further improve its effectiveness in clinical settings.

v. **Logistic Regression Classifier:**

23

Logistic Regression (LR) is a statistical machine learning model used for binary classification, making it able for predicting the likelihood of CVD. It estimates the probability of a condition based on predictor variables, offering a simple yet effective approach to early diagnosis. Studies in India and other countries have employed LR to improve CVD risk assessment. For example, [130] used regularized LR on the UCI dataset, achieving 90% accuracy, while Miao [131] enhanced LR with K-means++ and PCA for better data quality. Azis [132] applied 5-fold cross-validation, reporting accuracy between 80% and 88.29%, demonstrating the model's robustness. Additionally, Mahajan [133] found LR to be one of the most reliable models, achieving nearly 90% accuracy.

LR is favoured for its simplicity, ease of interpretation, and strong predictive capabilities, making it valuable in medical diagnostics[132],[130] . It can also integrate with methods like PCA to improve performance[131] . However, its reliance on linear relationships may limit accuracy in complex datasets, and it is sensitive to outliers[134]. Despite these drawbacks, LR remains a trusted tool in CVD prediction, with the potential for further enhancement through hybrid modelling approaches.

IV. **DATA SETS**

A dataset is a structured collection of data used to train machine learning models. In CVD prediction, key datasets include NIH's(National Institute of Health) multi-omics data for genetic risk assessment [135], UCI(University of California Irvine machine learning repository) Heart Disease for ML benchmarks[136] , IEEE Dataport(Institute of Electrical and Electronics Engineers) for ensemble models [137], UK(United Kingdom) Biobank for deep learning [138], and Kaggle for ML experimentation [139]. While these datasets enhance prediction, challenges like data quality and model interpretability require further validation[139][136]

Commented [U1]:

The NIH dataset, particularly from NHANES(National Health and Nutrition Examination Survey), has been instrumental in training [44] models for CVD prediction by providing extensive patient data, including demographics, medical history, and lifestyle factors[140]. It has also been used alongside CAC (collection of coronary artery calcium scans in AI-CVD studies to predict CVD events and mortality through AI-driven screening[141]. Researchers like Surya et al. applied neural networks and gradient boosting for early diagnosis, while M. Naghavi et al. [141], leveraged AI for enhanced risk assessments [140],[141]. Strengths of the dataset include its comprehensive variables and large sample size, which improve model reliability[142]. However, challenges such as limited population diversity and intensive data preprocessing require further validation for broader clinical application [143].

25 UCI Heart Disease dataset

The UCI Heart Disease dataset is widely used in ML research for CVD prediction, serving as a benchmark for testing various models. It includes diverse patient attributes, enabling the development of both traditional and advanced algorithms for early diagnosis. Studies have applied centralized and federated ML approaches, with SVM achieving 83.3% accuracy in a centralized setting and 73.8% in federated learning, enhancing privacy[142]. Regularized logistic regression demonstrated 90% accuracy, highlighting its predictive power [130], while CNN-based deep learning models outperformed traditional methods in efficiency and accuracy [144]. Researchers such as Rodriguez & Nafea explored model interpretability, Singh validated logistic regression, and Mansoor optimized deep learning techniques. The dataset's strengths include versatility across ML applications and support for privacy-preserving methods [145],[130]. However, its limited size and class imbalance may hinder model generalization[145], [146]. Addressing these challenges could enhance its real-world applicability.

The UK Biobank

The UK Biobank dataset, with data from over 500,000 participants, is vital for ML-based CVD prediction[147]. Researchers have used ML models to analyze genetic, lifestyle, and clinical factors. Nine models, including LSVM(Lagrangian Support Vector Machine), RF(Random Forest), and NN(Neural Network), utilized cardiac imaging and baseline characteristics [148]. ML tools like REGENIE and VariantSpark fine-mapped coronary artery disease loci[149] while survivalFM improved risk prediction by modeling factor interactions [150]. Strengths include comprehensive data and advanced ML techniques, enhancing sensitivity[149]. Challenges include complexity and computational demands, requiring solutions for better clinical integration [149].

IEEE Dataport

The IEEE Dataport dataset has been instrumental in improving ML models for CVD prediction, leveraging demographic and biometric data to identify high-risk individuals. Researchers have applied models like support vector machines, ensemble learning, and neural networks, achieving AUC values between 0.85 and 0.92[151]. Logistic regression, random forest, and K-nearest neighbors enhanced classification accuracy, while AI-integrated simulations reached 97% accuracy via a web-based tool[152]. Strengths include diverse data, high model accuracy, and user-friendly applications [151], [152]. However, demographic biases [153] and dataset size limitations [152] remain challenging. Expanding the dataset and addressing bias could further refine predictive models for clinical use.

Kaggle Dataset

The Kaggle dataset, comprising 70,000 clinical records, is widely used in ML-based CVD prediction[154]. Researchers have applied various models to improve accuracy and reliability. Ensemble learning combining Decision Tree, k-NN, and Naïve Bayes achieved over 99% accuracy [143], while deep neural networks outperformed traditional ML methods[155]. Feature selection using PCA with XGBoost enhanced model performance [156]. Random Forest demonstrated robustness with 88.04% accuracy [157], and SMOTE(Synthetic Minority Over-sampling Technique) addressed class imbalance[158]. Strengths include a large dataset and support for diverse ML techniques [143]. However, data imbalance and limited generalizability remain challenges [143],[158]. Validation is needed to improve real-world applicability.

V. TRAINING METHODS

Model training optimizes algorithms to recognize patterns and make accurate predictions. For cardiovascular disease (CVD) prediction, five key models—SVM, Logistic Regression, Random Forest, XGBoost, and KNN—rely on specialized training methods. SVM requires parameter tuning and feature scaling [159], [160], while Logistic Regression benefits from feature selection and regularization[161]. Random Forest improves accuracy through hyperparameter optimization and ensemble learning [159]. XGBoost enhances performance via gradient boosting, regularization, and pruning[159]. KNN depends on distance metric selection and feature scaling [159]. Computational efficiency varies, with ensemble models like Random Forest and XGBoost requiring more resources than Logistic Regression and KNN [159], [162]. Applying optimal training methods ensures effective CVD prediction.

A. Parameter Tuning and Feature Scaling

Hyperparameter tuning optimizes model settings to enhance performance, while feature scaling standardizes input data for balanced influence. In training SVM models for CVD prediction, GridSearchCV systematically explored parameter combinations, improving accuracy to 81.96%[163]. Pandu et al., also applied GridSearchCV and RandomizedSearchCV, though with a broader focus on ML models [164]. Kernel selection proved crucial, as Hoque et al. found polynomial SVM outperformed linear SVM[165]. Feature scaling, essential for distance-based calculations, was emphasized in preprocessing strategies [166]. While these techniques enhance precision, they can be computationally demanding and risk overfitting if misapplied (Pandu et al., 2023)[164].

B. Feature Selection and Regularization

Feature selection identifies the most relevant variables, while regularization prevents overfitting by penalizing large coefficients. In training logistic regression models for CVD prediction, Redhaei et al. applied Genetic Algorithm (GA), Particle Swarm Optimization , and Arithmetic Optimization Algorithm , achieving 98.59% accuracy[167]. [168]. Behki & Pal integrated PSO to refine feature selection, enhancing performance and reducing complexity [168]. Regularization techniques improved model generalization, with Singh et al. achieving 90% accuracy in CVD prediction[169]. Despite improving accuracy, these methods require careful parameter tuning and can reduce interpretability, affecting clinical adoption [161], [169].

C. Ensemble Learning and Hyperparameter Optimization

Ensemble learning, which combines multiple models for improved accuracy, and hyperparameter optimization, which fine-tunes parameters for optimal performance, have been effectively used in training Random Forest models for CVD prediction. Goad & Deore applied ensemble-based techniques, significantly boosting prediction accuracy [170]. Chithra et al. integrated a CNN-based Random Forest model, achieving 94.8% accuracy[171]. Bayesian optimization further enhanced model robustness on datasets like Cleveland and Framingham [172]. Despite improved accuracy and feature importance identification, challenges like computational cost, data dependency, and limited interpretability remain, necessitating explainable AI techniques such as SHAP for clinical adoption [173],[172]

D. Boosting, Regularization and Pruning methods

Boosting, a technique that enhances weak learners iteratively, is key to improving XGBoost's accuracy in

cardiovascular disease (CVD) prediction. Studies by [173]Zhou and[174] Harshit et al. show that XGBoost outperforms KNN and Naïve Bayes in accuracy and F1 scores, making it highly effective for medical data [175]. On the other hand Regularization, which penalizes model complexity to prevent overfitting, ensures XGBoost remains generalizable to unseen data, crucial in clinical applications [176]. The use of L1 (Lasso) and L2 (Ridge) regularization further enhances interpretability [175]. Pruning, which removes unimportant branches in decision trees, refines model simplicity and boosts reliability in clinical settings[177], [178]. Despite these strengths, hyperparameter tuning and high computational demands pose challenges, requiring optimized implementation to balance overfitting and underfitting for improved CVD prediction.

E. Distance Metric Selection and Feature Scaling Method

Distance metric selection, which determines how similarity is measured between data points, and feature scaling, which ensures uniform contribution of all features, are critical in optimizing KNN for cardiovascular disease prediction. Prasetyo et al. found that using Cosine distance with 7 neighbors achieved 90.761% accuracy, outperforming Euclidean distance[179]. Feature scaling and selection further enhanced performance, with Arif et al. demonstrating that Principal Component Analysis (PCA) improved accuracy from 83.6% to 90.16%[180], while Islam et al.used Information Gain and Correlation Heatmaps for dimensionality reduction[181]. These methods boost KNN's predictive power, but challenges remain, as performance depends on careful distance metric selection, dataset size, and feature distribution [182].

VI. EVALUATION OF MACHINE LEARNING MODELS

Model evaluation is the process of assessing a machine learning model's performance to ensure its reliability in real-world applications, particularly in cardiovascular disease (CVD) prediction. Key metrics include accuracy, AUC-ROC, sensitivity, specificity, and confusion matrix analysis. Accuracy measures correct classifications, with XGBoost and ensemble methods achieving high scores[183], [184]. AUC-ROC assesses class separation, where XGBoost and LightGBM excel [184]. [185] Sensitivity gauges CVD case detection, improved by Random Forest with SMOTE [186]. Specificity, reducing false positives, is strong in logistic regression and decision trees [187]. Confusion matrix analysis details classification outcomes, widely used in SVM and logistic regression[183]. Addressing class imbalances with SMOTE further improves model performance[186],[185]

A. Accuracy, Precision, Recall and F1 Score

XGBoost consistently delivers the highest performance in [37] prediction, reaching up to 98.5% accuracy, 99.1% precision, 98.3% recall, and a 98.7% F1 score, thanks to its powerful ensemble learning techniques [188], [184]. However, its computational demands can be a drawback[159]. Random Forest follows closely with an accuracy of 89%, offering strong precision and recall due to its robustness in handling complex data [189], [190]. Logistic Regression, while simpler, achieves 81–85% accuracy and is particularly effective at identifying true-positive cases [184], [190]. SVM performs moderately, with 74% accuracy, but its sensitivity to parameter tuning affects consistency [190]. KNN struggles the most, reaching only 69% accuracy, largely due to its reliance on proper feature scaling[190]. While XGBoost leads in predictive power, models like Random Forest and Logistic Regression provide a balance between accuracy and computational efficiency, making them viable options for real-world applications.

B. AUC-ROC

Random Forest stands out as the top performer in CVD prediction, achieving an AUC-ROC of 0.89, especially when combined with ensemble techniques like voting and stacking [191]. It also maintains high accuracy (88.04%), surpassing XGBoost, SVM, and KNN[192]. The use of SMOTE further strengthens its ability to handle class imbalances [186]. XGBoost follows closely with an AUC of 0.79 but is generally outperformed by Random Forest in predictive accuracy[157], [191]. SVM shows promise but does not exceed Random Forest in AUC-ROC performance, though resampling techniques like SMOTE improve its [189], [192] . KNN struggles in comparison, being highly sensitive to dataset size and feature selection, which limits its performance [189],[193]. Logistic Regression, often used as a baseline model, remains less effective than ensemble methods, even with LASSO regularization [186], [189], [193]. While Random Forest dominates in AUC-ROC performance, the best model choice depends on dataset characteristics and specific clinical needs.

C. Sensitivity

Random Forest and XGBoost show the highest sensitivity in CVD prediction, excelling in identifying true positives[186], [192]. Random Forest performs best with SMOTE for class imbalance [186], while XGBoost benefits from hyperparameter tuning[189]. Logistic Regression, though simpler, achieves good sensitivity with LASSO. SVM improves with SMOTE but lags behind ensemble models[186], [189]. KNN struggles, as performance depends on parameters and dataset quality[189], [193]. Model choice depends on data and interpretability needs.

D. Confusion Matrix

Random Forest consistently outperforms other models in CVD prediction, achieving high accuracy (88.04%) and excelling in sensitivity and specificity, especially when combined with SMOTE[186], [192]. XGBoost is a strong

contender, offering high accuracy with proper hyperparameter tuning but requiring more computational power[189], [192]. Logistic Regression, valued for its simplicity, serves as a reliable baseline but lags in accuracy and sensitivity [186],[189]. SVM improves with resampling techniques but remains less effective than ensemble models. KNN struggles with accuracy and is highly dependent on parameter tuning [189], [193]. While Random Forest leads overall, model choice depends on data constraints, computational resources, and interpretability needs.

3) In the evaluation, XGBoost is a top choice for CVD prediction due to its high accuracy, efficiency, and ability to handle complex data. Its ensemble learning captures nonlinear relationships better than models like Logistic Regression, SVM, and KNN, achieving ROC AUC scores up to 0.9973[194]. It manages correlated data like hospital-level effects [195] and integrates feature selection methods such as Grey Wolf Optimization to reduce overfitting [196]. Unlike CNNs, which struggle with sparse tabular data, XGBoost excels in medical datasets [197]. While its computational demands are higher, its predictive power makes it ideal for CVD risk assessment in clinical settings [198].

VI. CHALLENGES AND OPPORTUNITIES IN ML BASED CVD PREDICTION

A. Model Interpretability

Interpretability challenges hinder ML adoption for CVD prediction in Uganda. Complex models lack transparency, making clinical trust difficult[199] . Uganda's diverse healthcare data adds complexity, requiring advanced feature selection [200] Explainable AI techniques, like rule extraction and attention mechanisms, improve transparency[201] . Better interpretability fosters trust, adoption, and personalized care, essential for AI-driven CVD prediction.

B. Data Quality and Imbalance

This includes incomplete or skewed datasets bias models toward majority classes, reducing sensitivity for high-risk patients [202],[203]Traditional models struggle with imbalanced data, affecting accuracy and recall [204]. Resampling techniques like SMOTE and ensemble methods like boosting improve predictions [205],[204] Advanced solutions, including weighted Random Forests and GANs, further mitigate imbalance [202]. Fine-tuning XGBoost and standardized preprocessing enhance accuracy and generalizability [203],[206]. Adapting these strategies to Uganda's healthcare system is key to improving ML-driven CVD prediction and reducing mortality.

C. Computational and Infrastructure Limitation

ML-based CVD prediction in developing countries like Uganda faces challenges in data availability, computational resources, and system integration. Limited access to quality

datasets affects model training, requiring better data collection and integration [207]. Infrastructure constraints hinder the deployment of complex models, though cloud-based solutions can provide scalable resources. Advanced models like deep learning demand sophisticated implementation, which may be difficult in resource-limited settings, emphasizing the need for simpler, efficient alternatives [208]. Developing user-friendly interfaces can improve adoption in clinical workflows. Strategic partnerships and tailored ML approaches are essential to overcoming these barriers and enhancing CVD prediction in Uganda.

D. Generalization across Diverse population

ML-based CVD prediction faces challenges in generalizing across diverse populations, especially in Uganda and other developing countries, due to variations in data quality, risk factors, and inherent biases. Limited access to standardized datasets affects model accuracy when applied to different regions [206]. Population-specific factors, such as genetic markers and chronic inflammation in HIV-positive individuals, must be integrated to improve predictions [209]. Biases in sampling can lead to disparities, requiring techniques like resampling to enhance fairness[210] . Adaptable models using data optimization and ensemble learning perform better across diverse datasets [207]. Overcoming these challenges demands not just technical solutions but also a deeper understanding of local health profiles and socio-economic influences.

D. Opportunities for Advancing ML in CVD

Prediction

Advancing machine learning (ML) in cardiovascular disease (CVD) prediction in Uganda and developing countries presents a major opportunity to improve early detection and healthcare outcomes. ML models like Random Forest, Gradient Boosting, and Neural Networks have achieved up to 99% accuracy, enhancing risk assessment and timely interventions [211],[207]. Ensemble learning further boosts accuracy by capturing complex patient data patterns [209]. Leveraging diverse datasets, including those from the UCI repository, refines model performance [212], while feature engineering and hyperparameter tuning enhance predictive accuracy[207] . Integrating ML into cloud-connected healthcare systems ensures accessibility for clinicians, enabling better risk assessment and personalized treatments [213]. However, challenges like data privacy, healthcare integration, and region-specific risk factors—such as HIV-related CVD risks in Sub-Saharan Africa—must be addressed for effective implementation [209].

VII. FUTURE DIRECTIONS

A. Improving Data Quality

Improving data quality is crucial for advancing machine learning ML in CVD prediction in Uganda and other developing countries. Hybrid feature selection methods, such as the CVD Prediction Framework , enhance data reliability by filtering irrelevant features, improving model performance[214]. Data cleaning and standardization, as seen in Uganda's weather data processing, can be applied to CVD datasets to eliminate inconsistencies and improve accuracy [215], [216]. Health policies promoting risk reduction strategies—such as tobacco control and physical activity—can also lead to better data collection and quality[217]. Data optimization frameworks, including standardization and oversampling techniques, enhance model generalizability across populations [206]. However, challenges like limited data collection points and resource constraints persist, requiring collaboration between policymakers, healthcare providers, and technologists to ensure sustainable improvements in predictive accuracy.

B. Advancing Model Interpretability and Trust

This involves balancing accuracy through use of ensemble models like XGBoost, when integrated with user-friendly tools like Streamlit, improve clinical usability[208]. The Randelistic algorithm, combining multiple classifiers, enhances accuracy and reduces false negatives [218](Ananthi et al., 2023). Explainable AI (XAI) methods, such as SHAP and LIME, address the "black box" issue, fostering trust among clinicians [219], [220]. Exploring alternative algorithms and optimizing features can boost generalization and reduce bias, ensuring broader applicability. Future efforts should focus on expanding datasets, clinical trials, and cost-effective AI integration to enhance real-world adoption.

C. Deploying Scalable and Resource-Efficient Solutions

Addressing high computational demands in ML-based CVD prediction in Uganda and developing countries requires scalable, cost-effective solutions. Leveraging existing healthcare infrastructure, such as HIV/AIDS care systems, can support cardiovascular care by training specialists and establishing regional centres [57]. Mobile-based solutions like CVD Magic enable early risk detection without expensive lab tests, improving accessibility [221],[222]. Innovative approaches, including DNA computing, reduce computational load through parallel processing [221], while optimized deep learning models enhance prediction accuracy with minimal resources[223]. Grid computing consolidates available resources to boost efficiency in financially constrained settings [224]. Overcoming challenges in technical expertise, funding, and long-term sustainability is crucial for successful implementation.

VIII. CONCLUSION

This review underscores the potential of machine learning in cardiovascular disease prediction, highlighting both its advancements and challenges in diverse healthcare settings. Among various models, XGBoost, a powerful gradient boosting algorithm, emerges as the most effective due to its superior accuracy, robustness, and ability to handle imbalanced datasets [55], [189]. Its efficiency in feature selection, stability, and adaptability makes it well-suited for real-world clinical applications, particularly in resource-limited settings like Uganda. While challenges such as data quality, interpretability, and computational constraints persist, leveraging XGBoost with optimized preprocessing and explainability techniques presents a promising path toward more reliable and impactful CVD prediction models.

IX. ACKNOWLEDGEMENT

This review was conducted with the guidance and support of esteemed academics, I. Ssibimba Bashir, a graduate of Bachelor of Computer Science from Kampala International University (Class of 2024), extend my sincere appreciation to my lecturers, Mr. Azizi W. [12], a lecturer at Kampala International University and a PhD candidate in Computer Science, and Mr. Jamir Ssebadduka, a lecturer at the [12] International University of East Africa, Uganda, and a PhD candidate [2] Computer Science. Their valuable insights and expertise greatly contributed to the development of this work.

REFERENCES

- [1] R. Kukar and R. T. Lee, "Cardiovascular diseases," Drug Discovery Today: Disease Models, Accessed Feb. 05, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1573404914000001>
- [2] J. J. Joseph et al., "Comprehension & Management of Cardiovascular Risk Factors for Adults with Type 2 Diabetes," *Journal of the American Heart Association: Circulation*, vol. 145, no. 9, pp. 722–730, Jun. 2022, doi: 10.1161/CIR.0000000000004480/FILE/Supplemental
- [3] G. A. Roth et al., "Global burden of cardiovascular diseases and Risk Factors, 1990–2019: Update From the Global Burden of Disease Study of the Collaborative Network," *J. Am. Coll. Cardiol.*, vol. 76, no. 23, pp. 2882–2921, Dec. 2020, doi: 10.1016/j.jacc.2020.11.007.
- [4] S. V. Venneri et al., "Heart Disease and Stroke Statistics—2021 Update," *Circulation*, vol. 145, no. 8, pp. E254–E345, Feb. 2021, 10.1161/IR.00000000000000938.
- [5] G. A. Mensah et al., "Global Burden of Cardiovascular Diseases and Risks, 1990–2022," *J. Am. Coll. Cardiol.*, vol. 82, no. 26, pp. 2380–2475, Dec. 2023, doi: 10.1016/j.jacc.2023.11.007.
- [6] G. A. Mensah, G. A. Roth, and V. Fuster, "The Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019," *J. Am. Coll. Cardiol.*, vol. 76, no. 20, pp. 2329–2332, Nov. 2019, doi: 10.1016/j.jacc.2019.10.019.
- [7] A. Basalgia et al., "Ranking countries 2020: Hypertension control in Africa by 2025," *Cardiovasc. J. Afr.*, vol. 20, no. 4, pp. 262–273, Jul. 2017, doi: 10.5306/CVJA.2017.4001.
- [8] D. Adaroye et al., "Prevalence, awareness, treatment, and control of hypertension in Nigeria in 1995 and 2005: Findings from two nationally representative surveys," *J. Clin. Hypertens.*, vol. 13, no. 5, pp. 963–977, May 2001, doi: 10.1111/j.17414242.
- [9] M. D. Cooper et al., "World Heart Report 2023: Confronting the World's Number One Killer," *World Heart Fed.*, pp. 1–52, 2023. [Online]. Available: <https://world-heart-federation.org/wp-content/uploads/World-Heart-Report-2023.pdf>
- [10] V. Nagurny, F. Wounou, and F. Rhéaume, "Determinants of gross domestic savings in public and nongovernmental institutions (ANGI) approach to categorization," *J. Econ. Survey*, vol. 9, no. 1, pp. 1–19, Dec. 2020, doi: 10.1196/jes-000484-0029-17ARTICLES_7.
- [11] J. Kayembe, B. K. Wanyonyi, A. Katalinzi, B. Lontong, and F. Nsulu, "Hypertension in women, men, and children in Africa: A systematic review," *BMC Cardiovasc. Disord.*, vol. 15, no. 1, pp. 1–11, Aug. 2021, doi: 10.1186/s12871-021-0542-4/ARLEN.
- [12] A. Sana, G. Bhattacharya, A. Bhattacharya, and A. Mohanty, "Machine Learning Algorithm and Applications," *Arif. Intell. Eng. Appl.*, vol. 1, no. 1, pp. 1–12, Jun. 2023, doi: 10.12097/01003000001-1-MACHINE-LEARNING-ALGORITHMS-APPLICATIONS-ARADNA-SANA/GUARADA-BHATTACHARYA-AVISHEK-ATTRAJIT-TAMA-MEHERA.
- [13] Abubakar, S. Syamakudin, V. Mantalina, and S. Mdennou, "The role of mathematics in machine learning for disease prediction: In-depth review in the healthcare domain," *Adv. Statistcs Sci. Eng. Technol.*, vol. 6, no. 4, pp. 02494010–02494010, Sep. 2024, doi: 10.20877/ASSET_V604948.
- [14] S. M. Mun, M. S. Khan, M. Shover, and A. Karay, "Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL): A Comprehensive Overview on Techniques, Applications and Research Directions," *2nd Int. Conf. Recent Comput. Source Sys. (ICRCS-23) 2024-Proc.*, pp. 1404–1409, Jul. 2024, doi: 10.1109/CSCS49000.2024.9428168.
- [15] M. P. Howard, S. Lu, K. Koenig, A. Słobodzieniak, and H. C. Verkhotur, "Deep Learning Architectures," *Stat. Comput.*, vol. 30, pp. 1–54, Jun. 2020, doi: 10.1007/s00368-019-01056-0_1.
- [16] Amr El-Shafie, Pallavi Mehta, and S. Naseemullah, "Beyond traditional analytics: Exploring hybrid deep learning techniques for heart disease detection," *World Adv. Adv. Res.*, vol. 23, no. 1, pp. 31–32, Jul. 2024, doi: 10.9374/WAAR.2024.23.1.0226.
- [17] N. L. Rose, S. K. Mellick, O. Kaya, and J. Rose, "Emerging trends and future directions in machine learning and AI learning architectures," *Appl. Mech. Learn. DeepLearn. Archit. Tools*, Oct. 2024, doi: 10.3093/3935-81-08271-3.
- [18] H. Senthil, R. Tejaiah, P. Kumar Suresh, U. Giri, S. Kumar/Mohapatra, and N. Ramchandran Arun, "FUTURE-BASED IMPLEMENTATION OF MACHINE LEARNING ALGORITHMS FOR CARDIOVASCULAR DISEASE PREDICTION," *J. Mech. Geomat. Mat. Sci.*, vol. 19, no. 11, pp. 109–125, Nov. 2023, doi: 10.1080/17502489.2023.253020.
- [19] J. Wu et al., "Effect of behavior and risk of cardiovascular diseases and prognosis among individuals with cardiovascular diseases: a systematic review and meta-analysis of 7 prospective cohort studies," *Int. J. Hyg. Environ. Health*, vol. 21, no. 1, pp. 1–17, Dec. 2024, doi: 10.1163/S121966422401586-79/2024S2.
- [20] H. C. Kim et al., "2018 Korean Society of Hypertension guidelines for the management of hypertension: Part II Epidemiology of Hypertension," *Clin. Hypertens.*, vol. 25, no. 1, pp. 1–16, Aug. 2019, doi: 10.1186/s12873-019-0370-0.
- [21] V. Venneri, S. C. Ferreira, S. J. Hartigan, and R. Kothiyal, "A Random Forest Algorithm for Predicting Congestive Heart Failure Status with Learning Styles," *2020 Int. Conf. Future Eng. Eng. Technol. (FutEdu-Digit. Convex. Intell. Syst. Green Technol. Educ. Statistic. Community Dev. ICYEE 2023 - Proceedings*, pp. 162–164, Oct. 2023, doi: 10.1109/ICYEE5078.2023.9548199.
- [22] V. Ignatova, A. Sidorov, and S. Kekhov, "Random forests with parameterized ensembles-based in formation gains for classification and regression problems," *Precis.*, vol. 10, Jan. 2024, doi: 10.7717/PEERJ-CS.1775.
- [23] K. J. Taher, A. M. Alshabani, and D. A. Zuberi, "Data Mining Classification Algorithms for Analyzing Soil Data," *Asian J. Res. Geogr.*, vol. 1, pp. 17–28, May 2021, doi: 10.9734/ABCO5/2021-VMS30196.
- [24] Q. A. Montañez López, A. Montañez López, and J. Crossan, "RandomForest for Generic Prediction," *Machine Struct. Mach. Learn. Methods Generic Predict.*, pp. 653–661, Jan. 2022, doi: 10.1007/978-3-030-90010-0_45.
- [25] V. Xu and V. Nicu, "Data Mining Prediction Based on Support Vector Machine Model," *Highlights Sci. Eng. Technol.*, vol. 102, pp. 311–320, Jul. 2024, doi: 10.5409/YDAS/74.
- [26] "PDF Application of Support Vector Machines in Machine Learning (2024) | Mohammad Khalid Sallam Ma's Work" [Accessed: Feb. 10, 2025]. [Online]. Available: <https://paperkit-papers-of-support-vector-machines-machine-learning-qipqip.pdf>
- [27] R. Giorda, S. Ferri, D. Lenini, and D. Conforti, "An Overview on the Advancements of Support Vector Machines in Healthcare Applications: A Review," *Information*, vol. 15, no. 4, Apr. 2024, doi: 10.3390/INF15040255.
- [28] Kumar Bhakti B. Singh, Ashishna Sharma, Aditya Verma, Renu K. Marwah, and Dr. Vandita Bhagat, "Machine Learning for the Malaria Disease Prediction System," *Int. J. Sci. Res. Comp. Eng. Inf. Technol.*, vol. 10, no. 3, pp. 673–684, Jan. 2024, doi: 10.52628/CSEITAN.0217.
- [29] M. Shabid, A. Pugazh, Mr. Anujit Chakraborty, Mr. Ajithi Gopal, Mr. Hemant Nandkumar, and Mr. Gaurav Patwal, "Prediction of Multiple Diseases using Machine Learning Techniques," *Int. J. Adv. Res. Sci. Comput. Technol.*, pp. 206–214, Apr. 2024, doi: 10.40173/SARCT-7029.
- [30] R. Raman, D. Torelli, J. Yalçınay, A. Brown, and J. C. D. Cruz, "K-Nearest Neighbors in Cardiology: A Processing Tool for Heart Disease Prediction," *2023 IEEE 16th Int. Conf. Humanasi. Nanotechnology, Inf. Technol., Eng. and Manag. (HNCN 2023)*, Nov. 2023, doi: 10.1109/HNCN60074.2023.1059264.
- [31] S. Chaitanya and K. Zhu, "An Analysis of the K-Nearest Neighbor Classifier to Predict Benign and Malignant Breast Cancer Tumors," *J. Stakeholder Res.*, vol. 12, no. 4, Nov. 2023, doi: 10.4611/1B3RHS.V1204.5577.
- [32] V. R. and R. Ganguly, "Exploring the Power and Practical Applications of K-Nearest Neighbors (KNN) in Machine Learning," *J. Comput. Allied Biol.*, vol. 2, no. 1, pp. 8–15, Feb. 2024, doi: 10.6999/JCA.202402.
- [33] "Research on Disease Prediction and Health Supplement Recommendation Algorithms Based on KNN Algorithm" [Accessed: Feb. 10, 2025]. [Online]. Available: <https://paperkit-papers-of-knn-algorithm-k-nearest-neighbor-algorithm-for-disease-prediction-and-health-supplement-recommendation-algorithms-based-on-knn-algorithm.pdf>
- [34] S. V. Raghava, G. Zein-Nabihah, Z. Nabilah-Billah, A. S. Palivava, and H. Abd-Shakir, "K-Nearest Neighbor Algorithm for Heart Disease Detection: A Comparative Evaluation of Multiple and Multivariate Diseases," *6th Int. Soc. Res. Rev. Int. Conf. Statist. J. Stat. ISSRI 2023 - Proceedings*, pp. 237–241, Dec. 2023, doi: 10.1109/ISSRI60074.2023.10407899.
- [35] Y. Wu, "Heart Disease Prediction Using Gradient Boosting Decision Trees," *Proc. 1st Int. Conf. Eng. Manag. Inf. Technol.*, pp. 527–535, Jan. 2024, doi: 10.23010/202830000005.
- [36] Y. Guo, L. Huo, Y. Guo, M. An, and W. Xiong, "A housing-based prediction model for house price prediction," *J. Comput. Appl. Math.*, vol. 12, no. 2, pp. 15–19, Mar. 2024, doi: 10.1049/JCM.2308062.
- [37] I. Mohamed Hassan, "Boosting Learning Algorithms for Chronic Disease Prediction: A Review," *Int. J. Machine Learn. Comput.*, vol. 30, no. 2, pp. 22–30, Oct. 2024, doi: 10.25100/IJMLC.V3002.906.
- [38] M. H. Gakire, "The use of Machine Learning in Predicting Disease Progression," *NEJM REPORT Int. J. Res. Med. Sci.*, vol. 5, no. 1, pp. 1–4, Jan. 2024, doi: 10.9296/NEJMRS20485_5_1_005.
- [39] B. R. Wankar, N. V. Kulkarni, A. V. Jadhav, S. R. Bhowmik, and S. M. Kothiyal, "Comparative Deep Learning and Convolutional Neural Network Based Disease Prediction Using Convolutional Neural Networks for Early Detection," *Edu. Endocrinol. Train. Perioperative Health Technol.*, vol. 10, May 2024, doi: 10.4108/EETH.10.6919.
- [40] D. A. L. Kalath, N. N. Thunir, and S. S. Al-Hadidi, "Malaria Disease Prediction Based on Convolutional Neural Networks," *J. Appl. Eng. Technol. Sci.*, vol. 5, no. 2, pp. 1165–1181, Jun. 2024, doi: 10.37593/IAETS.V5E2.
- [41] P. P. T. Jagtap, "Disease Prediction Using CNN Algorithms," *Indian Sci. J. Res. Eng. Manag.*, vol. 08, no. 05, pp. 1–5, May 2024, doi: 10.5591/ISREMD.M20240527.
- [42] F. Shah, V. Rajesh, N. A. Gudali, and J. K. Dabi, "Deep CNN in Healthcare," *Deep Learn. Biomed. Signal. Mod. Anal.*, pp. 221–236, Aug. 2024, doi: 10.1209781/020205149-14.
- [43] A. Abdulla, S. Syamakudin, V. Mantalina, and S. Mdennou, "The role of mathematics in machine learning for disease prediction in deep learning in the healthcare domain," *Adv. Statistcs Sci. Eng. Technol.*, vol. 6, no. 4, pp. 02494010–02494010, Sep. 2024, doi: 10.20877/ASSET_V604948.
- [44] N. Shanmugam, K. Harish Kumar, T. Jayashree, and S. Kartik, "Disease Prediction Using Artificial Intelligence: Neural Networks," *J. Artif. Intel. Comput. Networks*, vol. 6, no. 1, pp. 1–14, Mar.

- [45] A. Bajaj, "Enhancing Healthcare Predictions with Deep Learning Models," *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 21, pp. 2729–2730, Mar. 2024, doi: 10.1609/AAAI.V38S21.30543.

[46] Y. Hu, P. Huang, W. Hong, Q. Li, L. Liu, and J. Tang, "In-Depth Insights into the Application of Recurrent Neural Networks (RNNs) in Traffic Prediction: A Comparative Review," *Algorithm*, vol. 11, no. 19, pp. 366–388, Sep. 2024, doi: 10.3390/ALGO11090366.

[47] S. T. Shanbhag and D. I. Shasha, "Optimized deep learning framework and unsupervised deep learning methods for disease prediction," *J. Med. Syst.*, vol. 45, no. 6, pp. 951–958, Jul. 2023, doi: 10.1007/s10916-023-02474-9.

[48] N. M. Mohamed-Ali and R. Vincent, "Deep belief network-based approach for detecting Alzheimer's disease using multi-class data," *Comput. Struct.*, *Biostruct.*, vol. 21, pp. 1681–1690, Feb. 2023, doi: 10.1016/j.csb.2023.02.021.

[49] N. M. Mohamed-Ali and R. Vincent, "Deep belief Network and Approach for Detecting Alzheimer's disease using the multi-class data," *Comput. Struct.*, *Biostruct.*, vol. 21, pp. 1681–1690, Feb. 2023, doi: 10.1016/j.csb.2023.02.021.

[50] "PDF | Bioconductor Disease Prediction Model Construction Based on Computer AI design learning Technology (2024) | Lin Yang | 1 Citation." Accessed: Feb. 10, 2023. [Online]. Available: <https://paperkit.org/research-on-disease-prediction-model-construction-based-on-computing-ai.html>

[51] H. Byeon et al., "Deep Neural Network Model for Enhancing Disease Prediction on an Auto Encoder based Deep Learning," *SLAM Technical*, vol. 29, no. 3, pp. 10145–10145, May 2024, doi: 10.1016/j.slam.2024.10145.

[52] M. Guttman, S. Dvorkin, S. Hettler, L. Palyor, and R. Bernstein, "Deep learning genome-wide genetics through bijective sparse layer-wise layers for disease genetic association and disease risk prediction," *bioRxiv*, vol. 2123, doi: 10.1101/3533.2023.536673.73.V1.

[53] L. S. Cia, E. Kurni, P. L. Stofford, M. Irwan, H. Yu, and K. Wong, "Identifying Disease of Interest With Deep Learning Using Clinical Data," *J. Clin. Anesth.*, vol. 17, no. 1, pp. 1–10, Mar. 2005, doi: 10.3136/JCKM.2005.3877.

[54] L. He, Q. Qu, D. Shi, S. Chang, and Y. Wang, "A robust real-world attentional graph neural network for mri-to-mri disease associations prediction," *Brief. Biostat.*, vol. 25, no. 6, Sep. 2024, doi: 10.1002/BIOSTAT.1584.

[55] N. R. Kumar, K. Sharma, and R. Harris, "A tree-based ensemble approach to disease prediction from patient records," *2023 IEEE 7th Conf. Comput. Technol. (CCT)*, pp. 1–5, Dec. 2023, doi: 10.1109/CCT5980.2023.9556.

[56] Z. Yang, A. Mitra, W. L. de Berlinhoff, and M. Wu, "Transferable 3D medical encoder-decoder model to enhance prediction of future outcomes using electronic health records," *Nat. Comput. Sci.*, vol. 14, no. 1, Mar. 2023, doi: 10.1088/1469-7688/54/2/0253-0257.

[57] H. Lin et al., "OneClassDF: Building an interpretable Transfer Learning Bidirectional Encoder Representations from Transfer via a Framework for Longitudinal Self-supervised Cancer Patients," *Med.*, vol. 2023, doi: 10.1007/s11510-023-01850-1.

[58] A. Take, F. Yamada, and S. Zaitsev, "Generalized multi-criteria binarization: Multi-criteria estimation of the most probable MHD MRI diagnosis for stroke patients," *PLoS One*, vol. 16, no. 10, October, Oct. 2021, doi: 10.1371/JOURNAL.PONE.0258855.

[59] Z. Chen, M. Balon, and M. Brown, "Boosting Transformer and language model: GridNet for initial decision in telemedicine," *Proc. J. Med. Image Comput. Graphics*, vol. 3, pp. 332–340, Jan. 2023, doi: 10.1101/6551.V1/2023.JAG-IND/NET/3.3.2.

[60] M. A. Shabot, "A brief look at domain learning for early stage diabetes risk prediction," *J. Supercomput.*, vol. 80, no. 10, pp. 1940–1948, Mar. 2024, doi: 10.1007/s11229-024-02611-2.

[61] S. K. Banerjee and C. Sharabi, "Clinical insight: Comparative Analysis of deep Learning Models for Disease prediction using Multi-classifiers," *Int. J. Eng. Comput. Eng.*, *Comput. Electron. (ICECE)*, vol. 2024, Apr. 2024, doi: 10.1109/ICECE50072.2024.958476.

[62] A. M. Morton, A. Seher, H. Mohammadi, K. S. R. Moaveni, M. H. Zafar, and T. S. ElBoraei, "Enhancing Cardiovascular Disease Prediction via Hybrid Deep Learning Architecture: A Step Towards Smart Healthcare," *2023 Solid State Conf. Eng. Proc. Electr. Control Telecommunications Eng. (EECTC23)*, pp. 1–6, Nov. 2023, doi: 10.1101/100073.2023.103160.

[63] P. S. and V. Malapati, "Healthcare monitoring and recommendation model: Hybrid AI deep learning Enabled Disease Diagnosis Framework," *Appl. Sci.*, vol. 12, no. 21, 10238, May 2022, doi: 10.3390/120210238.

[64] J. Han, Y. Xiong, Y. Feng, G. Xu, and Y. Liu, "Enhancing Disease Prediction with a Hybrid CNN-LSTM Model," *Comput. Theory Probl. Eng. Sci.*, vol. 4, no. 2, pp. 8–14, Feb. 2024, doi: 10.5130/TPES.2024.040202.

[65] "Mechanistic Design and Scaling of Hybrid Architectures," arXiv.org, vol. abs/2403.17644, Mar. 2024, doi: 10.48393/ARXIV.2403.17644.

[66] S. Yang, X. Su, K. Xu, Y. Liu, T. Xiong, and Z. Yang, "Hybrid Architecture-Based Evolutionary Relief Neural Network for Disease Prediction," *IEEE Trans. Evolut. Comput.*, vol. 8, no. 4, pp. 1–16, Jun. 2024, doi: 10.1109/TEVC.2024.3350001.

[67] E. Al-Khatib, "Learning-to-learn Learning Architecture for Efficient Emotion Recognition in Audio Processing," *J. Adv. Research Trends Comput. Sci.*, vol. 11, no. 10, pp. 135–140, Nov. 2023, doi: 10.1101/7472.2023.11.84075.

[68] "Hybrid Architectures as Deep Reinforcement Learning in Heterogeneous Wireless Networks," arXiv.org, vol. abs/2206.00009, Jun. 2022, doi: 10.48393/ARXIV.2206.00009.

[69] G. Arfeneva, "A Hybrid-CNN-ENN Model for Enhanced Asthma Diagnosis: A Comparative Study of Machine Learning and Deep Learning Techniques," *J. Appl. Artif. Intell. Med.*, vol. 7, no. 2, p. 306, May 2024, doi: 10.5140/IAIM.V7I2.2024.5989.

[70] D. Daswani, S. Parwan, N. Daliyath, U. E. Libharwala, S. Dahal, and S. Simonov, "COVID-19 Detection using Hybrid CNN-ENN Architecture with Transfer Learning from X-Rays," *Curr. Med. Imaging Rev.*, vol. 20, Aug. 2023, doi: 10.1741/5762.962006/2023.709737.

[71] L. Arampatzis, A. Gichoula, W. Liotta, T. Maguet, and M. Kiliaris, "A Hybrid Classification Model of Artificial Neural Networks and Non-linear Support Vector Machine," *Int. J. Data Sci. Anal.*, vol. 8, no. 2, pp. 47–47, Jan. 2022, doi: 10.11616/IJDSA.2022.98012.

[72] "Hybrid Architectures for Deep Reinforcement Learning in Heterogeneous Wireless Networks," arXiv.org, vol. abs/2206.00009, Jun. 2022, doi: 10.48393/ARXIV.2206.00009.

[73] G. Arfeneva, "A Hybrid-CNN-ENN Model for Enhanced Asthma Diagnosis: A Comparative Study of Machine Learning and Deep Learning Techniques," *J. Adv. Res. Artif. Intell. Med.*, vol. 7, no. 2, p. 306, May 2024, doi: 10.5140/IAIM.V7I2.2024.5989.

[74] D. Daswani, S. Parwan, N. Daliyath, U. E. Libharwala, S. Dahal, and S. Simonov, "COVID-19 Detection using Hybrid CNN-ENN Architecture with Transfer Learning from X-Rays," *Curr. Med. Imaging Rev.*, vol. 20, Aug. 2023, doi: 10.1741/5762.962006/2023.709737.

[75] L. Arampatzis, A. Gichoula, W. Liotta, T. Maguet, and M. Kiliaris, "A Hybrid Classification Model of Artificial Neural Networks and Non-linear Support Vector Machine," *Int. J. Data Sci. Anal.*, vol. 8, no. 2, pp. 47–47, Jan. 2022, doi: 10.11616/IJDSA.2022.98012.

[76] S. Singh, S. Arora, N. Srivastava, K. Meng, A. Wazir, and H. Adluru, "Smart Healthcare System for Disease Prediction using hybrid Machine Learning," *Brief. Adv. Med. Biol.*, vol. 2024, Feb. 2024, doi: 10.1109/BAMM5000.2023.985975.

[77] T. J. Johnson, A. D. Davis, D. P. Gitter, and D. Peter, "Reduced blood vessel segmentation from retinal fundus images using a hybrid CNN-ENN model," *Comput. Vis. Image Underst.*, vol. 199, Dec. 2019, doi: 10.1016/J.CVIU.2019.16203.

[78] S. Singh, S. Arora, N. Srivastava, K. Meng, A. Wazir, and H. Adluru, "Smart Healthcare System for Disease Prediction using hybrid Machine Learning," *Brief. Adv. Med. Biol.*, vol. 2024, Feb. 2024, doi: 10.1109/BAMM5000.2023.985975.

[79] X. Dong, Z. Wu, Y. Wan, C. Shi, and Q. Ma, "A review on learning ensemble," *Front. Comput. Sci.*, vol. 16, no. 2, pp. 241–258, Apr. 2024, doi: 10.1007/s11704-019-0200-Z.

[80] P. Egorov, B. Gorji, K. Umugulu, P. Salimzadeh, E. Savenkov, and S. M. Kumar, "Program and prediction of the COVID-19 pandemic using machine learning," *Neurosci. Neuroeng. Review*, vol. 25, 2023, Dec. 2023, doi: 10.1007/s11704-019-0160-0.

[81] L. Wang, C. Zhang, and L. Li, "A Hybrid CNN-Transformer Model for Predicting N-Stage and Starvation in Newborns and Early Childhood Patients," *Sci. Technol. Geograph.*, vol. 10, no. 10, pp. 1676–1693, Oct. 2024, doi: 10.31975/OMGOGA/2024/101023.

[82] Z. Xiaoyu, "An Efficient Hybrid CNN-Transformer with Frequency Bias Bridging for Medical Image Segmentation," *Front. Bioengineering Biotech.*, vol. 12, Apr. 2024, doi: 10.31975/OMGOGA/2024/101023.

[83] J. W. Kim, A. Umez Khan, and I. Bantoci, "Systematic Review of Hybrid Vision Transformer Architectures for Diagnostic Image Analysis," *Front. Artif. Intell. Med.*, vol. 13, no. 10, pp. 1001–1009, Oct. 2024, doi: 10.31975/OMGOGA/2024/101023.

[84] S. Norilisa, P. Basit, and Z. Zeale, "Explaining Artificial Intelligence in Diagnosis of Images in Natural Networks and Artificial Neurons: A review of Decoding Interpretation," *IEEE PES Insur. Smart Grid Technol. Conf.*, Oct. 2023, doi: 10.1109/ISROPEP7670.2023.10489265.

[85] P. Sharif, S. Arora, A. Kalia, and V. Rajkumar, "Explaining Artificial Intelligence for Cybersecurity," *Intell. Comput. Cybersecurity*, vol. 19, no. 1, pp. 1–10, Oct. 2023, doi: 10.1002/9781911610904.ch1.

[86] Z. Pan and P. Mohan, "Explaining Artificial Intelligence," *Explain. AI Cybersecurity*, pp. 29–51, Jan. 2023, doi: 10.1007/s11704-023-0447-9.

[87] M. S. Aqib, A. Ullah, and D. Aslam, "Explaining Machine Learning Model for Chronic Kidney Disease Predictor," *Advances*, vol. 15, no. 10, pp. 443–455, Oct. 2024, doi: 10.31975/OMGOGA/2024/101023.

[88] S. Abbas, A. Qasim, M. Faruq, M. Saleem, M. Ahmad, and M. Arshad, "Smart Vaccine Temperature Logistics Online Prediction Using Ensemble Artificial Intelligence," *Sensor. Servic.*, vol. 24, no. 20, pp. 661–681, Oct. 2024, doi: 10.3390/S2-2306011.

[89] "Explainable-Diagnosis Prediction through Near-Symbolic Integration," *Proc. AAAI*, vol. 38, 2024, doi: 10.16093/ARXIV.2410.01855.

[90] V. N. Manz, N. Agarwal, and K. Krishna Somayaji, "Decision Tree-Based Explainable AI for Diagnosis of Diabetes Mellitus," *Proc. Conf. Comput. Appl. Ind. Eng. (CAIE)*, vol. 2024, pp. 947–952, Aug. 2024, doi: 10.1109/ICBEC50001.2024.937020.

[91] "An Interpretable Rule-Matrix Model for Hybrid-Biased Models based on Surrogate Trees – Skirtos," *AI*, vol. 2024, doi: 10.48393/ARXIV.2407.202007.

[92] J. Matthes, C. Clunie, and S. Kohli, "Interpreting Explainable AI: Explainable Artificial Intelligence (XAI) for the Diagnosis of Disease Prediction Models in Medical Rehabilitation Systems," *2024 IEEE Conf. Inf. Comput. Technol. Approach. Health. Monit. Sens. Syst.*, *Proc. IEEE*, vol. 2024, pp. 1–5, Mar. 2024, doi: 10.1109/SAHM50001.2024.9301932.

[93] G. Atman and S. Kalrahan, "Comparing Decision Trees and Association Rules for Stock Market Expectations in BIST 100 and SET 100," *Sci. Eng. Res.*, vol. 69, no. 3, pp. 459–475, 2022, doi: 10.47747/SAER.2022.0024.

[94] R. T. "Explainable AI: Developing Interpretable Deep Learning Models for Medical Diagnosis," *Int. J. Artif. Intell. Explic.*, vol. 6, no. 4, pp. 1–10, Jul. 2024, doi: 10.31975/OMGOGA/2024/101023.

[95] M. Kumar, K. Ramamurthy, and H. Garg, "Explainable AI: Disease Detection with Reasoning," *Proc. Adv. Conf. Appl. Artif. Intell. Comput. Eng. CAIE*, vol. 2024, pp. 222–227, Jan. 2024, doi: 10.1109/CAIE50001.2024.9301933.

[96] I. D. Mysore and B. Jeo, "An Explorable Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction," *Information*, vol. 15, no. 7, pp. 394–394, Jul. 2024, doi: 10.31960/INFSO.2024.9301934.

[97] N. Shukla, A. Jadhav, and V. Prakharpatay, "A Novel Explainable AI Model for Medical Data Analysis Using Ensemble Learning," *Comput. Res.*, vol. 14, no. 12, pp. 121–137, Mar. 2024, doi: 10.2478/CR-2024-0007.

[98] C. Mutha-Valsaraju, J. A. Qureshi, D. Omritkar, and B. Bhattacharya, "Explainability tools for disease diagnosis: What are they, why should we trust them?" *Proc. IEEE*, vol. 2024, doi: 10.1109/PRR50001.2024.9301935.

[99] L. P. Doria and A. de Weil, "A Taxonomy of Explainable Bayesian Networks," *Comput. Comput. Inf. Sci.*, vol. 18, pp. 1342–1350, Feb. 2024, doi: 10.1007/s11704-024-0611-5.

[100] "A Framework for Multi-modal Learning: Jointly Modeling Inter- & Intra-Modality Dependencies," *MoDA24*, doi: 10.48393/ARXIV.2409.2024.59315.

[101] "The Evolution of Multimodal Model Architecture," *MoDA24*, doi: 10.48393/ARXIV.2405.17927.

[102] P. Li, Z. Li, and J. Guo, "An Explorable Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction," *Information*, vol. 15, no. 7, pp. 394–394, Jul. 2024, doi: 10.31960/INFSO.2024.9301935.

[103] Y. Okuda, M. Akayama, and Y. Sakakihara, "Deep learning of multimodal networks with topological regularization for drug dispensing," *J. Cheminform.*, vol. 16, no. 1, pp. 1–10, Jan. 2024, doi: 10.1145/3597403.2024.10001.

[104] S. Li and R. Zhang, "A novel multi-stage deep capsule spectral convolutional network with multi-scale feature fusion for medical image classification," *Neurosci. Neuroeng. Review*, vol. 17, pp. 1085–1086, Oct. 2024, doi: 10.1007/s11704-024-01000-0.

[105] P. Liang, S. Guo, and Z. Peng, "Foundations in Trends in Methodology: Machine Learning Principles, Challenges, and Open Questions," *Comput. Surv.*, vol. 56, no. 10, Apr. 2024, doi: 10.1145/3597403.

[106] Y. Okuda, M. Akayama, and Y. Sakakihara, "Deep learning of multimodal networks with topological regularization for drug dispensing," *J. Cheminform.*, vol. 16, no. 1, pp. 1–10, Jan. 2024, doi: 10.1145/3597403.2024.10001.

[107] S. Li and R. Zhang, "A novel multi-stage deep capsule spectral convolutional network with multi-scale feature fusion for medical image classification," *Neurosci. Neuroeng. Review*, vol. 17, pp. 1085–1086, Oct. 2024, doi: 10.1007/s11704-024-01000-0.

[108] L. Wang, "A hybrid multimodal graph for disease prediction," *Brief. Bioinform.*, vol. 25, no. 3, Jul. 2024, doi: 10.1003/BIB.58435.

[109] R. Wanger, "A hybrid multimodal graph for disease prediction in disease prediction," *Methodol.*, vol. 29, no. 4, Apr. 2024, doi: 10.1611/YMETH24-00003.

[110] L. S. Ranganath, F. Shah, A. Bhagat, and L. S. Komanduri, "Heart Disease Detection Using Graph Neural Networks (GNNs)," *Proc. Conf. Recent. Comput. Syst. Comput. Syst. ICRCSE*, vol. 2024, pp. 136–144, Apr. 2024, doi: 10.1109/ICRCSE50001.2024.9301936.

[111] S. Venkateswaran, "Design of an Improved Model for Lung X-ray Disease Prediction Using Graph Neural Networks," *Neurosci. Neuroeng. Review*, vol. 17, pp. 665–668, Oct. 2024, doi: 10.1007/s11704-024-01000-0.

[112] W. Wang, "Heart disease prediction using the learning gridcode," *Theor. Neuro. Sci.*, vol. 51, no. 1, pp. 9–17, Oct. 2024, doi: 10.542/5475-5803/2024/10001.

[113] S. Yash, A. Yadav, S. D. Srivastava, and P. Rai, "Heart Disease Prediction Using Machine Learning," *Advanc. Sci. Eng. Eng. Manag.*, vol. 00, no. 000, pp. 1–4, Sep. 2024, doi: 10.31941/SEEM50001.

[114] L. S., "Predicting Heart Disease through Machine Learning Methods," *J. Integr. Sci. Techol.*, vol. 1, no. 1, pp. 1–10, Jan. 2024, doi: 10.31975/OMGOGA/2024/101023.

- pp. 829-842. Sep. 2024, doi: 10.31241/JISRT/MSRT24/SE9/82.
- [111] J. Basurto, "Heart Disease Prediction Using Machine Learning Algorithms," *J. Assoc. Comput. Emerg. Resour.*, vol. 4, no. 2, Oct. 2024, doi: 10.5635/JACER.4.2.2.V4.2.145.
- [112] Z. Almasi, "Abstract 41 (304): Discovering novel biomarkers and predicting cardiovascular disease using machine learning for cardiovascular diseases," *Circulation*, vol. 150, no. Suppl_1, Nov. 2024, doi: 10.1161/CIRCULATIONAHA.123.441684.
- [113] "Abstract 110 (304): Heart Disease Prediction: A Machine Learning Approach For Early Detection And Risk Assessment," Oct. 2024, doi: 10.48593/FAKVN.24/011478.
- [114] R. Gaurav, S. Fornari, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machines in Cardiac Applications: A Review," *Information*, vol. 15, no. 4, Apr. 2024, doi: 10.3390/INF15040233.
- [115] K. Srivastava, R. E. Pehlivan, C. Baldwin, and C. V. Y. Kumar, "S. Srivastava, R. E. Pehlivan, and C. Baldwin, "Cardiovascular Disease Prediction Using Machine Learning Models: A Review," *Int. J. Conf. Innov. Mach. Appl. ICMIA23 - Proc.*, pp. 630-667, Dec. 2023, doi: 10.1109/ICMIA40077.2023.9404542.
- [116] S. Tuncer, D. Demirci, and Y. Cakiroglu, "Analysis and Enhancement of Prediction of Cardiovascular Disease Diseases Using Machine Learning Models SVM, SGD, and XGBoost," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 1, Jan. 2024, doi: 10.14859/IJACSA.2024.013040.
- [117] V. Sen, H. Ran, and M. Venkatesh Karan, "Novel approaches for predicting risk factors of cardiovascular disease," *IEEE Trans. Reliab.*, vol. 17, no. 1, pp. 135-151, Mar. 2023, doi: 10.1109/TRIEE.2022.3222727.
- [118] Bahmanpour, Hosseini, Bilezikci, Arsan Dehghani, S. M. Saeid Hosseini, and Naseri Ben Shahr, "Heart Disease Prediction Using SVM," *J. Sci. Res. Arch.*, vol. 11, no. 2, pp. 187-193, Oct. 2024, doi: 10.30574/JSRRA.2024.11.2.0035.
- [119] D. Mahanta and A. Saha, "Predicting Cardiovascular Disease with Machine Learning: A Comparative Analysis of Classification Models," Jul. 2024, doi: 10.21038/S.8-429736.3.V1.
- [120] R. Venkateswaran, S. Ananthanarayanan, and S. Ganeshanaran, "Multi-parametric based adaptive SVM-based prediction for heart disease," *Int. J. Computer Commun. Control.*, vol. 18, no. 5, Aug. 2023, doi: 10.1583/700CC.2023.5.4994.
- [121] T. Dang, "Evaluating Cardiovascular Risk Predictors: Development of an Advanced XGBoost Model with Hospital-Level Random Forests," *Biostatistics*, vol. 11, no. 10, pp. 1039-1050, Oct. 2024, doi: 10.1093/BIOSTATISTICS/KZAB002.
- [122] Y. Zhou, "Prediction and Analysis of Cardiovascular Diseases Based on MLBoost," *2024 IEEE 2nd Conf. Data Privacy Comput. Appl. (COPCA) 2024*, pp. 1877-1881, Jun. 2024, doi: 10.1109/COPCA55300.2024.9600000.
- [123] D. Bhattacharya, Y. Singh, & A. Saha, "Advancing Cardiovascular Health Risk Assessment: A Comparative Study on Heart Attack Prediction Using kNN and Multinomial Cross-Validation," *2024 Int. Conf. Adv. Soft Eng. Sci. Eng. Technol. (ACREST) 2024*, pp. 1-5, Sep. 2024, doi: 10.1109/ACREST502024.2024.9744313.
- [124] C. Yan, Y. Xiong, S. Liu, E. Wang, & Y. Wang, "Machine Learning Models for Cardiovascular Disease Prediction: A Comparative Study," *J. Electr. Syst.*, vol. 20, no. 3, pp. 663-672, Apr. 2024, doi: 10.5273/JES.135.
- [125] S. N. Ant, A. M. Sepehri, S. Saeedi, and A. R. Jonaiia, "K-means/Fuzzy C-means Clustering and Multidimensional Machine Learning for Nearest Neighbors (KNN) and Support Vector Machine (SVM) Classification," *J. medical bioinform.*, vol. 5, no. 3, p. 1617, Jul. 2024, doi: 10.30806/MIB.V03.1644.
- [126] S. Kalra, A. Sharpen, and T. S. Durrani, "SVM Analytics for Heart Disease Using Machine Learning," *Int. J. Adv. Comput. Inf. Technol.*, vol. 20, no. 2, May 2024, doi: 10.1109/IACT.0156.24.4109024.
- [127] Samir K. Patel, "Random Forest-Based Prediction Models for Assessing Cardiovascular Disease Risk: The Impact of Age and Genetic Factors," *J. Electr. Syst.*, vol. 20, no. 3, pp. 663-672, Apr. 2024, doi: 10.5273/JES.135.
- [128] V. R. Barmpakis, V. Dattu, Mantri, and N. Vora, "Personalized Cardiovascular Disease Risk Stratification Using Random Forest: An Optimal Approach," *2023 IEEE World Conf. Appl. Inf. Serv. Comput.*, pp. 226-232, Mar. 2023, doi: 10.1109/WCASIC53763.2023.9649013.
- [129] K. B. Radford, "Optimizing Cardiovascular Disease Prediction: Harnessing Rankers for an Algorithm with Advanced Feature Selection," *Jn. 2024*, vol. 10, 21035.95, 38.5-384/2024/01V1.
- [130] S. P. Singh, A. Singh, and M. Kumar, "Beyond Traditional Methods Using Regularized Logistic Regression for Acute Myocardial Infarction Prediction," *Proc. - 2024 Int. Conf. Energy Adv. Comput. INSCOMP 2024*, pp. 247-251, Mar. 2024, doi: 10.1109/INSCOMP50224.2024.9620048.
- [131] H. Min, "Logistic regression for cardiovascular disease prediction by integrating PCA and Lasso++," *Proc. Natl. Acad. Sci.*, vol. 38, no. 1, pp. 126-132, Jun. 2024, doi: 10.1073/PNAS.23.1009536.NI23.10246913.
- [132] H. Ans, "Assessing the Performance of Logistic Regression in Heart Disease Detection through 5-Fold Cross-Validation," *J. Artif. Intell. Med. Infor.*, vol. 2, no. 1, pp. 1-11, May 2024, doi: 10.3679/JAIM.2023.10137.
- [133] V. Mahajan, E. Arshad, D. Sowmy, and A. Skypal, "Advanced Machine Learning for Risk Assessment of Cardiovascular Disease," *Int. J. Artif. Intell. Comput. Infor. Syst.*, pp. 1-6, Nov. 2023, doi: 10.1007/S00593-023-04003-1.
- [134] S. Wang, "Research on the heart attack prediction based on logistic regression," *Highlights Sci. Eng. Technol.*, vol. 65, pp. 155-158, Aug. 2023, doi: 10.3097/HSET.2023.1157.
- [135] W. Darroudi et al., "Multimodal AIML for discovering novel biomarkers and predicting disease using multi-level profiles of patients with cardiovascular diseases," *Proc. Sci. researc.*, vol. 14, no. 1, p. 2653, Nov. 2024, doi: 10.1088/2419-5093/24/1/015003.
- [136] Riaz Nasar, "Machine and deep learning techniques for cardiovascular disease detection," *J. Assoc. Comput. Emerg. Resour.*, vol. 4, no. 2, Oct. 2024, doi: 10.5635/JACER.4.2.2.V4.2.131.
- [137] N. Ghoshroy, Ramaiah, Adwajon, W. Malaran, and A. Albus Gradi, "Prediction of Cardiovascular Disease (CVD) in the Elderly Using Tree-based Ensemble Model," *2024 12th Int. Conf. on Adv. Comput. Inf. Technol. (ICACIT) 2024*, pp. 210-216, Aug. 2024, doi: 10.1109/ICACIT4641.2024.9309830.
- [138] R. Bhuvaneswaran, S. Karthikam, and T. Selvamurugan, "CNNET - Cardiovascular Disease Detection using Deep Convolutional Neural Network," *2024 Int. Data Privacy Conf. Technol. (DPCIT) 2024*, pp. 1-6, Apr. 2024, doi: 10.1109/DPCIT502024.2024.1057300.
- [139] N. V. DEOGA SAJAI, SIVA VARA PRASAD RAJU, and P. N. DEVI, "All-Around Medical Imaging and Heart Disease Diagnosis: A Deep Learning Approach for Automated Analysis and Enhanced Prediction Using Transfer Learning and Ensemble Model," *Int. J. Sci. Eng. Syst. JASSN*, vol. 10, no. 1, pp. 210-220, Oct. 2024, doi: 10.60607/JASSN.GV01.V01.242.
- [140] B. Sarav, V. Venkatesh, S. Vijayakumar, A. H. Narayanan, and R. Syed, "AI for Early CVD Diagnosis and Personalized Care," *Int. J. Med. Syst.*, vol. 11, no. 02, pp. 09-18, Jan. 2024, doi: 10.4787/70REMED2024.V4/01021008.
- [141] M. Naghavi et al., "Abstract #4403: AI4CV: Artificial Intelligence-Enabled Opportunistic Screening of Coronary Artery Calcium-Computed Tomography Scans for Predicting CVD Events and All-Cause Mortality: The Multi-Ethnic Study of Atherosclerosis (ME-SA)," *Circulation*, vol. 150, no. Suppl_1, Nov. 2024, doi: 10.1161/CIRCULATIONAHA.123.44067.
- [142] S. Shah et al., "Application of Machine Learning for Cardiovascular Disease Risk Prediction," *Congr. Statis. Resour.*, vol. 2023, no. 1, pp. 1-12, Mar. 2024, doi: 10.1158/2023.940666.
- [143] I. Islam, "Enhancing Cardiovascular Disease Prediction Accuracy through an Ensemble Machine Learning Approach," *Int. J. Conf. Inf. Model. Med. Infor.*, vol. 2, no. 2, pp. 95-100, Nov. 2024, doi: 10.3679/JIMI.2023.100215.
- [144] C. Minoche, S. E. Chait, and H. Naderi, "Advancing Heart Disease Prediction通过集成机器学习和迁移学习技术," *Proc. - 3rd Int. Conf. Adv. Geogr. Comput. Appl. Informatics. (CAGI) 2023*, May 2024, doi: 10.1109/ACAGI50104.2024.960247.
- [145] "Continued and Enhanced Heart Disease Classification Models Using UCI Dataset and their Shapley-value Based Interpretability," Aug. 2024, doi: 10.48593/FAKVN.24/0120435.
- [146] M. Sanagota, S. A. Karim, S. P. Bhattacharya, P. K. Giri, and P. B. Prakash Reddy, "Heart Disease Prediction Model Using MLR," *Int. J. Adv. Sci. Res. Technol.*, vol. 20, no. 3, Mar. 2024, doi: 10.3812/43870.18824/JSRTE.2024.0120416.
- [147] C. Stalder et al., "UK Biobank: an open access resource for understanding the causes of a wide range of complex diseases of middle and old age," *PLoS Med.*, vol. 12, no. 3, Mar. 2015, doi: 10.1371/JOURNAL.PME.1011779.
- [148] "Machine Learning Models for the Identification of Cardiovascular Diseases Using UK Biobank Data," *arXiv.org*, vol. 2407.07121, Jul. 2024, doi: 10.48593/ARXIV.2407.16721.
- [149] L. M. F. Stoye et al., "From Mapping Coronary Artery Disease Lesions in the UK Biobank Whole Genome Sequencing Data Optimizing Usage of the Clinical Based Results in Analysis Platform," *medRxiv*, Sep. 2024, doi: 10.1101/2024.09.23.2410018.
- [150] H. Jukkola and J. Rousu, "Machine learning for comprehensive interaction modelling improves disease risk prediction in UK Biobank," *medRxiv*, Aug. 2024, doi: 10.1101/2024.08.07.2401604.
- [151] A. Di Belotti and F. K. Alipourha and O. Olszak, "Predictive Modelling for Cardiovascular Disease in Patients Based on Demographic and Biometric Data," *ABUAD J. Eng. Res. Dev.*, vol. 7, no. 1, pp. 153-159, Apr. 2024, doi: 10.5092/AJER.D.2024.701.7015.
- [152] M. A. Sofi et al., "Enhancing Clinical Validation for Early Cardiovascular Disease Prediction through Streamlined Machine Learning," *Int. J. Biostat.*, vol. 14, no. 12, pp. 1305-1308, Jan. 2024, doi: 10.1309/BIOSTAT.K4121108.
- [153] F. L. L. Zhao, P. Wu, H. H. Qiu, W. Wei, and F. Petrone, "Abstract 14590: Evaluating Methodology Metrics for Machine Learning-based Cardiovascular Risk Model," *Circulation*, vol. 140, no. Suppl_1, Nov. 2023, doi: 10.1161/CIRCULATIONAHA.140.115589.
- [154] S. Shah et al., "Application of Machine Learning for Cardiovascular Disease Risk Prediction," *Congr. Statis. Resour.*, vol. 2023, no. 1, p. 943066, Jun. 2023, doi: 10.1158/2023.940666.
- [155] J. Li, "Prediction of cardiovascular disease based on machine learning," *Appl. Geospat. Eng.*, vol. 46, no. 1, pp. 60-66, Mar. 2024, doi: 10.5424/2758-2724.46.2024.014090.
- [156] K. Ahsan, M. S. R. Salim, B. S. Piray, M. I. Rahman, and M. Shahzad, "Multi-Focused Approach to Cardiovascular Risk Assessment by Utilizing Predictive Machine Learning and Clinical Data in a Unified Web Platform," *IEEE Access*, pp. 1-1, Jun. 2024, doi: 10.1109/ACCESS.2024.3406700.
- [157] K. Nihal, U. H. Seti, B. F. Ahmed, M. Arun, J. Rana, and V. G. Krishnan, "Cardiovascular Disease Prediction Through Machine Learning Algorithms," *Proc. 3rd Conf. Appl. Afr. Inf. Geogr. (CAIG) 2023*, May 2024, doi: 10.1101/19CAIG.0022.2024.1057499.
- [158] R. Parashar, V. Datt, A. Karim, and N. Iftekhar, "Boosting the Accuracy of Coronary artery Disease Prediction Using Machine Learning," *2023 Int. Conf. Adv. Inf. Technol. (TICIT) 2023*, pp. 1-6, Jun. 2023, doi: 10.1009/ICAIT5759.2625.1058370.
- [159] R. G. - et al., "Review on Heart Disease Prediction using Machine Learning Approaches," *Int. J. Microbiol. Infect.*, vol. 20, no. 3, Oct. 2024, doi: 10.50968/1DMR.2024.01602945.
- [160] D. Mallawa et al., "Predicting Cardiovascular Disease with Machine Learning: A comparative Analysis of Classification Models," Jul. 2024, doi: 10.1103/19CAIG.0022.2024.1057495.V1.
- [161] A. M. Neelakant, A. A. Lagakas, A. Olszak, and D. D. Olszak, "Logistic regression technique for coronary artery disease prediction," *Fibrodys. J. Sci.*, vol. 8, no. 4, pp. 266-275, Aug. 2024, doi: 10.3300/JFD.2024.008024.
- [162] V. S. Behera and D. Bala, "Implementing Stacking with Cross-validation for Heart Disease Prediction," *Int. J. Physiol. Qual. Assur.*, vol. 15, no. 10, pp. 1493-1496, Sep. 2024, doi: 10.2425/19PQQA.15.3.
- [163] D. P. Molena, H. K. Ogata, G. Smith, and B. Bui, "Optimizing Heart Disease Prediction Model with GridSearchCV for Hyperparameter Tuning," *2023 Int. Conf. Geogr. Geoinf. Unqiquitous Geosp. IC-GUG 2023*, pp. 1-6, Mar. 2024, doi: 10.1109/IC-GUG500782.2023.1038772.
- [164] E. P. Cynthia, B. S. Mohamed Samir, W. S. Li, E. Imamura, J. A. Bryant, and M. J. Schmid, "Improved Machine Learning Algorithm for Heart Disease Prediction Based on Hyperparameter Tuning," *Int. IEEE Inv. Comp. Appl. Eng. Technol. (ICECAT) 2023*, pp. 176-181, Sep. 2023, doi: 10.1109/ICECAT50041.2023.
- [165] Bahmanpour, Hosseini, Bilezikci, Arsan Dehghani, S. M. Saeid Hosseini, and Naseri Ben Shahr, "Heart Disease Prediction using SVM," *Int. J. Sci. Res. Arch.*, vol. 11, no. 2, pp. 412-420, Mar. 2024, doi: 10.3812/43870.18824/JSRTE.2024.0120455.
- [166] S. Radford, W. G. Kumar, D. M. Rau, M. V. V. K. Kartikey, and M. Arun, "Heart Disease Prediction Using GradientBoostCV and Random Forest," *EAI End-to-end Trans. Perservate Health Technol.*, vol. 10, Mar. 2024, doi: 10.4010/EETPT.10.552.
- [167] A. A. Radwan, W. A. Faris, and M. A. Atar, "Using Optimization Techniques in Feature Selection for Efficient Cardiovascular Disease Prediction," *2023 4th Int. Conf. Inf. Technol. - CIT*, 2023, pp. 1-4, Dec. 2023, doi: 10.1109/CIT588.2023.1043377.
- [168] P. Behlend and R. Pal, "Prediction of Heart Disease by Feature Selection Technique using Particle Swarm Optimization-based Logistic Regression," *2023 2nd Int. Conf. Feature Inf. Technol. INCFT 2023*, pp. 1-9, Nov. 2023, doi: 10.1109/INCFT5007.2023.1052450.
- [169] S. P. Singh, A. Singh, and M. Kumar, "Beyond Traditional Methods: Utilizing Regularized Logistic Regression for Heart Disease Prediction," *Proc. - 2024 Int. Conf. Adv. Comput. Inf. Technol. (INSCOMP) 2024*, pp. 249-251, May 2024, doi: 10.1109/INSCOMP50224.2024.10040048.
- [170] P. M. Gould and P. J. Ross, "Predicting and Analyzing Cardiovascular Disease via a Machine Learning Approach," *Int. J. Artif. Intell. Med. Infor.*, vol. 4, no. 3, pp. 153-165, Sep. 2024, doi: 10.5429/19PMT2410.
- [171] V. S. Chait, H. S. Shekhar, and D. M. Shabot Patel, "Heart Disease Detection Using Ensemble Machine Learning Technique Based on Rule Bases," *2nd IEEE Int. Conf. Netw. Multimed. Inf. Technol. (SMINT) 2024*, pp. 1-3, Aug. 2024, doi: 10.1109/MTCON5024.2024.10066991.
- [172] I. D. Misra and N. Jar, "Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction," *Information*, vol. 15, no. 7, pp. 394-394, Jul. 2024, doi: 10.3390/INF15070394.
- [173] S. Zhu, "Generalization Investigation in Heart Disease Prediction: Comparative Analysis of MLP," *medRxiv*.

- Dec. 2019, doi: 10.1007/s10430-019-1633-28.
- [74] S. Singh, S. Singh, A. Naray, S. Kanti, M. A. Wani, and A. Akhter, "Smart Healthcare System for Disease Prediction Using Hybrid Machine Learning," *Proc. - 2021 Int. Conf. Aspirer. Academ. Sustain. Syst. ICASSS* 2021, pp. 362–369, Aug. 2021, doi: 10.1109/ICASSS58487.2021.9262857.
- [75] "Autodesk Revit," Jan. 2014, doi: 10.405.90/ASME.V.1409.7461.
- [76] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Fuzzy. Comput. Sci.*, vol. 14, no. 2, pp. 241–260, Sep. 2020, doi: 10.1007/s11787-019-02268-z.
- [77] P. Gupta, B. Gopta, K. Krapushy, P. Solanki, E. Sreenivasan, and S. M. Kumar, "Prognosis and prediction of disease using hybrid machine learning framework," *Neurocomputing. Advances*, vol. 2023, Jan. 2023, doi: 10.1016/j.neuad.2023.011684.
- [78] L. Wang, C. Zhang, and L. Li, "A Hybrid CNN-Transformer Model for Predicting N-Stage and Survival in Lung Cancer," *Int. J. Comp. Theory. Technol.*, vol. 10, no. 10, pp. 1676–1687, Oct. 2024, doi: 10.3999/IJCTT.2024.1003547.
- [79] Z. Xian et al., "Enhancing Hybrid CNN-Transformer via Frequency-based Bridging for Medical Image Segmentation," *Proc. - 2024 Int. Conf. on CS-Soft. Technoplogy*, pp. 1–4, May 2024, doi: 10.1109/SSFT50730.2024.1003547.
- [80] J. W. Kim, A. Ueng Khan, and I. Bouziane, "Systematic Review of Hybrid Vision Transformer Architectures for Radiological Image Analysis," Jan. 2024, doi: 10.1101/20240124.12409265.
- [81] S. Naikar et al., "Transformers and large language models in healthcare: A review," *Artif. Intel. Med.*, vol. 151, pp. 102050–102050, Aug. 2024, doi: 10.1016/j.artmed.2024.102050.
- [82] D. Barua, P. Barua, and P. Ma, "Explaining Artificial Intelligence by Definition of Inputs in Neural Networks and Ensemble Forecasting of Electricity Consumption," *IEEE PES. Power. Syst. Gen. Distrib. Conf. Exh.*, Oct. 2023, doi: 10.1109/PWGUS50965964.2023.11048289.
- [83] F. Klaudius et al., "Explanation artificial intelligence in pathology," *J. Pathol. Anal.*, vol. 45, no. 2, pp. 135–140, Jun. 2024, doi: 10.1007/s00292-024-01676-7.
- [84] P. Stavroulakis, K. Astheri, A. Kotsiris, and V. Rupakheti, "Explorable Artificial Intelligence for Cybersecurity," *Hot. Comput. Cybersec.*, pp. 149–174, Oct. 2024, doi: 10.1002/9781119616919.CHT.
- [85] Z. Pan and P. Melis, "Explaining Artificial Intelligence," *Explor. AI Cybernetics*, pp. 29–51, Jan. 2021, doi: 10.1107/0796-3401-464799.2.
- [86] M. S. Al-Ali, A. Rehman, and D. Asif, "Explaining Machine Learning Model for Chronic Kidney Disease Prediction," *Algorithms*, vol. 17, no. 10, pp. 443–443, Oct. 2024, doi: 10.3390/ALGO1710045.
- [87] S. Akbar, A. Qasim, M. S. Imran, M. Salam, M. Ahmad, and M. A. Khan, "Smart Vision Transparency Efficient Online Disease Prediction Model Using Explainable Artificial Intelligence," *Sensors*, vol. 24, no. 20, pp. 6018–6018, Oct. 2024, doi: 10.3390/S24206018.
- [88] "Explaining Diagnosis Prediction through Neuro-symbolic Integration," *Proc. - 2024 IEEE*, doi: 10.48550/ARXIV.2010.01838.
- [89] V. N. Marpa, N. Apurva, and K. Krishnan Srivastava, "Distance Tree-Based Explainable AI for Diagnosis of Chronic Kidney Disease," *Proc. - 2024 Int. Conf. Adv. Comput. Appl. ICACA* 2024, pp. 947–952, Aug. 2023, doi: 10.1109/ICACA5908.2023.96122074.
- [90] "An Interpretable Rule Creation Method for Block-Basis Models based on Surrogate Trees – Skirtos," *Int. J. Comput. Inf. Syst. Eng.*, vol. 2024, doi: 10.48550/ARXIV.2047.2010.
- [91] J. Matheus, R. Chima, C. Stephan, and R. S. Kirby, "Interpretable Explainable Artificial Intelligence (XAI) for the Development of Product and Information Recommendation Systems," *Int. J. Adv. Inf. Technol. Approaches. Technol. Manag. Syst. Knowl.*, vol. 2024, pp. 1–5, Mar. 2024, doi: 10.1007/978-3-7091-360-0-2_0204.
- [92] G. Atman and S. Kahraman, "Comparing Decision Tree and Association Rules for Stock Market Expectations in BIST100 and BIST500," *Ser. Adv. Econ. Bus.*, vol. 69, no. 3, pp. 499–515, 2022, doi: 10.47733/SEAB.2022.224458.
- [93] R. T., "Explaining AI: Developing Interpretable Black-Box Learning Models for Medical Diagnosis," *Int. J. Med. Inf. Res.*, vol. 6, no. 4, Jul. 2024, doi: 10.36948/IMIR.2024.0004.20261.
- [94] M. Kumar, K. Korrapati, and H. Geng, "Explaining AI: Disease Detection with Reasoning," *Proc. - 2024 Int. Conf. on Big Data. Machine Learning. Deep Learning. Big Data. ML. DL. BDL. 2024*, doi: 10.1109/ICBDL4022.2024.10575155.
- [95] J. D. Matheus and N. Azev, "Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction," *Information*, vol. 15, no. 7, pp. 394–394, Jul. 2024, doi: 10.3390/INF15070394.
- [96] N. Sudhakarova, P. Sharad, and V. Prakash, "A Novel Explainable AI Model for Medical Data Analysis," *Int. J. Artif. Intel. Soft. Comput. Res.*, vol. 14, no. 2, pp. 121–137, Mar. 2024, doi: 10.24793/JASC.18.2024.4007.
- [97] C. S. Matute-Vallencia, J. A. Oquendo, D. Orozco-Beltran, and X. Barber, "Bayesian networks for disease diagnosis: What are they, who uses them and how? (Review)," Apr. 2025, doi: 10.2196/34995.48570.
- [98] L. P. de la Torre and de la Wal, "A Taxonomy of Explainable Bayesian Networks," *Comput. Comput. Inf. Sci.*, vol. 1342, pp. 220–235, Feb. 2024, doi: 10.1007/978-3-030-6013.9.14.
- [99] "A Framework for Multimodal Learning: Joint Modeling Inter- & Intra-Modality Dependencies," May 2024, doi: 10.48550/ARXIV.1706.13.
- [100] "The Evolution of Multimodal Model Architectures," May 2024, doi: 10.48550/ARXIV.2405.1972.
- [101] P. P. Liang, A. Zakhid, S. P. Merenyi, "Foundations Across Traditional Multimodal Machine Learning, Deep Learning, and Open Questions," *AIChE Comput. Surv.*, vol. 56, no. 1, Apr. 2024, doi: 10.1002/ACCS.6830.
- [102] V. Ohnaki, M. Akimoto, and Y. Sakakibara, "Deep learning of multi-modal networks with topological regularization for drug quantification," *J. Chemometr.*, vol. 16, no. 1, Aug. 2024, doi: 10.1106/SI.3221-0289-0097-Y.
- [103] S. Li and R. Zheng, "A novel interactive deep capsule spectral graph convolutional network with multi-relational graph for disease prediction," *Neural Networks*, vol. 175, pp. 106285–106285, Apr. 2024, doi: 10.1016/J.NEUNET.2024.106285.
- [104] Y. Zhou et al., "Multimodal functional deep learning for metabolomics data," *Brief. Bioinform.*, vol. 25, no. 5, Jul. 2024, doi: 10.1093/BIB/BBA484.
- [105] R. Wang et al., "Hybrid multimodal fusion for prob learning in disease prediction," *Med. Model.*, vol. 229, pp. 41–48, Sep. 2024, doi: 10.1016/j.YMED.2024.06.003.
- [106] L. S. Antropiotti, F. Shukla, A. S. Bhagat, and S. Komnenović, "Heart Disease Detection Using Graph Neural Networks (GNNs)," *Int. J. Comp. Electron. Syst. Comput. Syst. ICESCS* 2024 – Proc., pp. 1456–1461, Aug. 2024, doi: 10.1109/ICESCS50052.2024.10091813.
- [107] Sandeep Sharma, "Design of Improved Model for Lung Skin Disease Prediction Using Graph Neural Networks and Multiple Model Data Fusion," *Devel. J.*, vol. 27, no. 2, pp. 663–685, Aug. 2024, doi: 10.1007/s10430-023-00423-1.
- [108] W. Wang et al., "Heart disease prediction using grid line integration," *Theor. Nat. Sci.*, vol. 51, no. 1, pp. 9–15, Oct. 2024, doi: 10.38254/275-3805-51-2024-010144.
- [109] S. Yadav, A. Yadav, D. S. Srivastava, and P. Rai, "Heart Disease Prediction Using Machine Learning," *Asian Sci. J. Sci. Eng. Manage.*, vol. 08, no. 008, pp. 1–4, Sep. 2024, doi: 10.59041/ASJM/7904.
- [110] L. S. "Predicting Heart Disease through Machine Learning Methods," *Int. J. Innov. Sci. Res. Technol.*, pp. 829–842, Sep. 2024, doi: 10.3024/18181838724-0829892.
- [111] J. Hasan, "Heart Disease Prediction Using Machine Learning Algorithm," *J. Innov. Comput. Eng. Technol.*, vol. 2, Oct. 2024, doi: 10.53536/JICET.0218143.
- [112] Z. Ahmed, "Abstract 1410347: Discovering non linear features and predicting cardiovascular disease using AI/ML techniques for precision medicine," *Circulation*, vol. 159, no. Suppl. 1, Nov. 2024, doi: 10.1161/CIRCULATIONAHA.1510347.1410347.
- [113] "Advancements in Heart Disease Prediction: A Machine Learning Approach For Early Detection And Risk Assessment," Oct. 2024, doi: 10.48550/ARXIV.2410.14738.
- [114] R. Gruber, S. Ferreira, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Information*, vol. 15, no. 4, Apr. 2024, doi: 10.3390/INF15040235.
- [115] K. Srinivas, R. B. Palani, C. Balakrishnan, U. V. R. Kumar, R. Sugunan, and M. Anisa, "Support Vector Machine: Unraveling the Power and Versatility of SVM in Modern Machine Learning," *Int. J. Comp. Inf. Eng.*, vol. 15, no. 4, Sep. 2024, pp. 639–647, doi: 10.3997/1364-580X.102023-02340245.
- [116] S. Turner, D. Demirkilic, and V. Chandra, "Analysis and Enhancement of Prediction of Cardiovascular Disease Diagnosis using Machine Learning Models (SVM, SGD, and XGBoost)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 4, pp. 469–479, Oct. 2024, doi: 10.14896/IDACSA.2024.019349.
- [117] V. Seet Han Rau and M. Nordin Kurmi, "Novel approaches for predicting risk factors of atherosclerosis," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 1, pp. 183–189, 2013, doi: 10.1109/JBHI.2012.222771.
- [118] Ralented Hoque, Masum Billah, Ami Debbartha, S. M. S. Sankar Hossein, and Nurur Bin Shuru, "Heart Disease Prediction using SVM," *Int. J. Sci. Res. Tech.*, vol. 11, no. 2, pp. 421–424, Mar. 2024, doi: 10.3979/BEIJINGA.2024.11-2408.
- [119] D. Mallavarapu and V. Suresh, "Predicting Cardiovascular Disease with Machine Learning: A Comparative Analysis of Classification Models," Jul. 2024, doi: 10.22913/35.8.29780-351.
- [120] R. Verma, S. S. Venkateswaran, and S. Gundredda, "Multi-grained heart nodules SVM-based prediction model for heart disease," *Int. J. Geogr. Comput. Geod.*, vol. 18, no. 8, Aug. 2023, doi: 10.1080/13653390.2023.22494.
- [121] T. Bang et al., "Enriching Cardiovascular Risk Prediction: Development of an Advanced Xgboost Model with Hospital-Level Random Effects," *Biostatistics*, vol. 11, no. 1, pp. 1039–1050, Oct. 2024, doi: 10.3997/BEIJINGA.1101039.
- [122] Y. Zhou, "Prediction and Analysis of Cardiac vascular Disease Based on XGBoost," *2024 IEEE 2nd Int'l Conf. Image-Proces. Comput. Appl. ICIPCA* 2024, pp. 1877–1881, Jun. 2024, doi: 10.1109/ICIPCA5193.2024.10079218.
- [123] D. Bhambhani, V. Singh, R. Kumar, and L. Sharma, "Advancing Cardiovascular Health Risk Assessment: A Comprehensive Study on Heart Attack Prediction Using XGBoost Model and Cross-Validation," *2023 Int'l Conf. on Comp. Eng. Comput. Technol. (ICCECT) 2023*, pp. 1–5, Sep. 2024, doi: 10.1109/ICCECT52023.00541.
- [124] C. Yau, Y. Xiong, S. Li, L. Guo, and J. Wang, "Machine Learning Models for Cardiovascular Disease Prediction: A Comparative Study," Jun. 2024, doi: 10.1017/2024.05.27.594902.
- [125] S. N. Anisa, A. M. S. Sankar, S. Rani, and R. V. Rama, "Non-linear Parallel Learning Iterating Marginalized Multi-class Machine Learning k-Nearest Neighbors (k-NN) and Support Vector Machine (SVM)," *J. Multicoll. Inf. Commun.*, vol. 8, no. 1, 167, Jul. 2024, doi: 10.3986/MVI.1084784.
- [126] S. Kali, A. Srikanthreddy, A. M. Yaser, and S. Dharmar, "Predictive Analytics for Heart Diseases Using Machine Learning," *2024 Int'l Conf. Eng. Technol. INCE2024*, May 2024, doi: 10.1109/INCE2024.36156241093243.
- [127] Suranjan Karayannidis, "Random Forest-Based Prediction Models for Assessing Cardiovascular Disease Risk: Integrating Clinical and Genetic Factors," *J. Biostat. Syst.*, vol. 20, no. 3, pp. 663–672, Apr. 2024, doi: 10.23832/S273831522407204.
- [128] V. R. Brangash, D. Mani, Manu, and N. Vysa, "Personalized Cardiovascular Disease Risk Prediction Using Random Forest: An Optimized Approach," *Proc. - 2023 IEEE World Conf. Appl. Intel. Comput. ACC 2023*, pp. 226–232, Jul. 2023, doi: 10.1109/ACCION702023.2023.1004015.
- [129] K. Bi and R. A., "Optimizing Cardiovascular Disease Prediction: Harnessing Random Forest Algorithm with Advanced Feature Selection," Jun. 2024, doi: 10.21203/3.RS-3034700.V1.
- [130] S. P. Singh, A. Singh, and M. Kumar, "Beyond Traditional Methods: Utilizing Regularized Logistic Regression for Accurate Heart Disease Prediction," *Proc. - 2024 Int'l Conf. Zhenjiang Adv. Comput. INNZC-COMPAZ 2024*, pp. 247–253, May 2024, doi: 10.1007/978-3-031-60154-2_25.
- [131] H. Mu, "Logistic regression for cardiovascular disease prediction by integrating PCA and k-means++," *Int. J. Comp. Inf. Eng.*, vol. 35, no. 1, pp. 126–132, Jun. 2024, doi: 10.3986/MVI.1084784.
- [132] H. Attia, "Assessing the Performance of logistic Regression in Heart Disease Detection through 5-fold Cross Validation," *Int. J. Artif. Intel. Soft. Comput. Res.*, vol. 2, no. 1, pp. 1–11, May 2024, doi: 10.24793/JASC.18.2024.40072311.
- [133] Y. Maitra, E. Arshab, D. Soman, and L. Lohith, "Advanced Machine Learning for Risk Assessment of Coronary Artery Disease," *Int. J. Artif. Intel. Soft. Comput. Res.*, vol. 2025, pp. 1–6, Nov. 2023, doi: 10.1007/ICCECT999.12025.1042103.
- [134] S. Wang, "Logistic regression for heart attack prediction based on logistic regression," *Highlights Sci. Eng. Technol.*, vol. 05, pp. 153–158, Aug. 2023, doi: 10.4490/HSET.2023.1157.
- [135] W. DeGroot et al., "Multimodal AI/ML for discovering novel biomarkers and predicting disease using multi-omics profiles of patients with cardiovascular diseases," *Diseases*, vol. 14, no. 1, p. 2603, Nov. 2024, doi: 10.3390/DISEASES140102603.
- [136] Ram Naresh, "Machine and deep learning techniques for cardiovascular disease detection," *J. Innov. Comput. Eng. Technol.*, vol. 2, Oct. 2024, doi: 10.565/SOCECT-V42-131.
- [137] N. Ghosh and A. Banerjee, "Advances in Machine Learning for Risk Assessment of Coronary Artery Disease (CAD) in the Upcoming Years Using Tree-Based Ensemble Models," *2024 Int'l Conf. Comp. Eng. Comput. Technol. ICCECT 2024*, pp. 210–216, Aug. 2024, doi: 10.1109/ICCECT5164.2024.10079010.
- [138] R. Bhattacharyya, S. Karbigyan, and T. Salimuzzaman, "CVNET - Cardiovascular Disease Detection using Convolutional Neural Network (CNNs)," *Int. J. Comp. Electron. Syst. Comput. Syst. ICESCS* 2024 – Proc., pp. 1–6, Jul. 2024, doi: 10.1109/ICESCS2024.2024.10079006.
- [139] N. V. DURGA SAI SRIVI, VARA PRAKASH RAJU, and P. N. DEVI, "An Advanced Machine Learning-based Heart Disease Diagnosis: A Deep Learning Approach for Automated Analysis and Extraction of Predictors Using Ensemble Classifiers," *J. Artif. Intel. Gen. Syst.*, vol. 15, no. 1, pp. 210–229, Oct.

- [2024] doi: 10.1007/s10807-023-01601-242.

[140] B. Saray, B. Karahan, S. Yavuz, M. Akter, A. H. Nurayannan, and R. Syed, "AI for Early CVD Diagnosis and Prognosis," *J. Adv. Comput. Inf. Eng. Res.*, vol. 01, no. 02, pp. 99–109, Jan 2024, doi: 10.47673/IJREME2024/0100100002.

[141] M. Naghavi et al., "A Systematic Review and Meta-analysis of Machine Learning Approaches for Predicting Cardiovascular Disease Risk and All-Cause Mortality: The Global Study of Attributable Mortality (MS-EASy)," *Circulation*, vol. 150, no. Suppl_1, Nov. 2024, doi: 10.1161/CIRCULATIONAHA.121.MS-EASy.1444883.

[142] S. Shahid et al., "Application of Machine Learning for Cardiovascular Disease Risk Prediction," *Comput. Biol. Neurosci.*, vol. 2023, no. 1, pp. 1–12, Mar 2023, doi: 10.1155/2023/948690.

[143] I. Ilkhan, "Enhancing Cardiovascular Disease Prediction Accuracy through an Ensemble Machine Learning Approach," *Jn. Artif. Intell. Med.*, vol. 2, no. 2, pp. 95–105, Jun 2024, doi: 10.36703/JAIM2024/02125.

[144] C. Manousi, S. K. Chatterjee, and H. Nalani, "Advancing Heart Disease Prediction through Synergistic Integration of Machine Learning and Deep Learning Techniques," *Proc. - 3rd Int'l Conf. Adv. Comput. Comput. Appl. Informatics. ACCI-2024*, Dec. 2024, doi: 10.1007/978-981-996160-1-0006.

[145] "Centralized and Federated Heart Disease Classification Models Using UCI Dataset and Four Shirley-value Based Interpretability," *Aug 2024*, doi: 10.5201/ARXIV.2408.0163.

[146] M. Sangathee, S. A. Kumar, P. B. Bharath, K. Parva, and P. B. Prakash Reddy, "Heart Disease Prediction Using ML," *J. Inst. J. Sci. Res. Technol.*, pp. 2019–2033, Apr 2024, doi: 10.30124/IJSRST1884-2244.20240100.

[147] C. Sudhoff et al., "UK biobank an open access resource for identifying the causes of a wide range of complex diseases of middle and old age," *PLoS Med.*, vol. 12, no. 3, Mar 2015, doi: 10.1371/journal.pmed.1007187.

[148] "Machine Learning Models for Identification of Cardiovascular Diseases Using UK Biobank Data," *arXiv*, vol. org, arXiv:2407.16721, Jul 2024, doi: 10.4236/ALSP.20240716721.

[149] L. M. F. Soe et al., "Fine-Tuning Convolutionary Autoencoder Loss in the UK Biobank Whole Genome Sequencing Data by Optimizing the Cloud-Based Research Analysis Platform," *bioRxiv*, Sep 2024, doi: 10.1101/240857; arXiv:2409.23120.

[150] H. Almazroo and J. Rizzo, "Machine learning for cardiovascular disease prediction: improving vs risk prediction in the UK Biobank," *bioRxiv*, Aug 2024, doi: 10.1101/2024.08.24.554104.

[151] A. D. Bahlani, K. F. Alengundi, and M. Olfekar, "Predictive Modeling for Cardiovascular Disease in Patients With Type 2 Diabetes Mellitus," *Am. J. Engg. Appl. Eng.*, vol. 2, no. 1, pp. 153–159, Apr 2024, doi: 10.3192/AJEER.2024/010115.

[152] M. A. Safar et al., "Ensemble Learning Methods for Early Cardiovascular Disease Prediction through Simulation, AI and Web Technologies," *Diagnóstica*, vol. 14, no. 12, pp. 1308–1308, Jun 2024, doi: 10.3924/DIAGNOSTIC18412108.

[153] F. Li, J. Zhao, S. Wu, H. Wang, Q. Wan, and J. Patterson, "Abstract 15589: Evaluation Methods to Mitigate the Bias in Machine Learning-Based Cardiovascular Risk Model," *Circulation*, vol. 140, no. Sept 1, Nov 2022, doi: 10.1161/CIRB.140.30919; p. 15589.

[154] S. Shahid et al., "Application of Machine Learning for Cardiovascular Disease Risk Prediction," *Comput. Biol. Neurosci.*, vol. 2023, no. 1, pp. 94/00–94/06, Jan 2023, doi: 10.1155/2023/94006.

[155] J. Li, "Prediction of cardiovascular disease based on machine learning," *Appl. Geogr.*, vol. 46, no. 1, pp. 60–66, Mar 2024, doi: 10.1016/j.apgeog.2024.02.010.

[156] K. Akbar, M. R. S. Khalil, B. H. S. M. Ibrahim, M. M. Ibrahim, and M. Shafiqul, "Multi-Faceted Approach to Cardiovascular Risk Assessment by Utilizing Predictive Machine Learning and Clinical Data in a Unified WebPlatform," *IEEEXAS*, pp. 1–10, Jun 2024, doi: 10.1109/IEEEXAS42024.9406023.

[157] K. Nihalani, U. B. Sofi, B. F. Ahmad, M. Arjan, J. Ravi, and V. G. Krishnan, "Cardiovascular Disease Prediction Through Machine Learning Algorithms," *Proc. - 3rd Int'l Conf. Adv. Comput. Appl. Comput. ICCAC-2024*, pp. 385–388, Jun 2024, doi: 10.11000/ALSP.002222/0024.0007499.

[158] R. Parangon, D. A. Kumar, and N. Bhati, "Investigating the Accuracy of Cardiovascular Disease Prediction Through SMOTE," *2023 3rd Int'l Conf. Comput. Engg. Comput. Technol. ICCECT-2023*, pp. 1–6, Jun 2023, doi: 10.1109/ICCECT57733.2023.9540203.

[159] R. G. et al., "Review on Heart Disease Prediction using Machine Learning Approaches," *Int. J. MobiComp. Res.*, vol. 6, no. 4, Oct 2024, doi: 10.3948/IJMR.2024/080103.

[160] D. Mahanta and A. Saha, "Predicting Cardiovascular Disease with Machine Learning: A Comparative Analysis of Classification Models," *Int. J. Modelling. Optim.*, vol. 2024, doi: 10.17616/8ZK8-3X85-3Z62.

[161] A. M. Nwosu, A. A. Lujade, A. Olibe, and D. O. Obahene, "Logistic regression technique for cardiovascular disease prediction," *Facult. J. Sci.*, vol. 8, no. 4, pp. 266–275, Aug 2024, doi: 10.3303/FJS-1804-0024.

[162] N. V. Praveen and D. Balaji, "Implementing Stacking with Cross – Validation for Heart Disease Prediction," *Int. J. Comput. Appl.*, vol. 15, no. 05, pp. 149–153, Sep 2014, doi: 10.22375/IJCAI-15_3_02.

[163] D. P. Maitra, R. K. Gupta, C. Smith, and R. Basu, "Optimizing Heart Disease Prediction Model with Gradient-Descent (or Hyperparameter Tuning)," *J. Int'l Conf. Geogr. Oper. Optimiz. Comput. IC-GOOC-2024*, p. 1–6, Mar 2024, doi: 10.1109/ICGOOC8708.2024.970272.

[164] E. P. Cytus, B. M. Mohamed-Salem, S. L. Li, E. Lamine, A. Alfrangi, and M. A. Aljaleel, "Improved Machine Learning Algorithm for Heart Disease Prediction Based on Hyperparameter Tuning," *Natl. REIEEE Conf. Artif. Intell. Engg. Technol. ICET-2025*, pp. 176–181, Sep 2025, doi: 10.1109/ICET4941.2025.9210753.

[165] R. Rahman, Hognar, Mansur Bilal, Arif Dehmani, S. M. Sadiq Hussain, and Nomir Bin Shumr, "Heart Disease Prediction Using SVM," *J. Sci. Res. Arch.*, vol. 11, no. 2, pp. 412–420, Mar 2024, doi: 10.3074/SRAS-1204-112-0485.

[166] S. Radhika, G. K. Kumar, D. M. Rao, M. V. V. P. Kurnagi, and M. Anil, "Heart Disease Prediction Using Deep Learning Model," *EduLearned Int'l Conf. Persv. Health. Technol.*, vol. 10, Mar 2024, doi: 10.1403/EPHELT.2023.

[167] A. Al-Rashed, W. A. Faris, and M. A. Al-Boraei, "Utilizing Optimization Techniques as Feature Selection for Effective Cardiovascular Disease Prediction," *2022 2nd Int'l Conf. Adv. Inf. Technol. ACIT-2022*, pp. 1–6, Dec 2022, doi: 10.1007/978-981-996160-3-0022.

[168] P. Behdi and B. Pal, "Prediction of Heart Disease by Feature Selection Techniques using Particle Swarm Optimization based on Logistic Regression," *2023 2nd Int'l Conf. Artif. Intell. Engg. Technol. ICET-2023*, pp. 1–9, Nov 2023, doi: 10.1109/ICET50073.2023.9245265.

[169] S. P. Singh, A. Singh, and A. Kumar, "Bayesian Trial Methodology Utilizing Regularized Logistic Regression for Accurate Heart Disease Forecasting," *Proc. - 2nd Int'l Conf. Econ. Infor. Adv. Comput. INCOMAC-2024*, pp. 247–251, Jul 2024, doi: 10.1109/INCOMAC50024.2024.940048.

[170] P. M. Gund and P. J. Deane, "Predicting and Analyzing Cardiovascular Disease through Ensemble Learning Approaches," *Int. J. MobiComp. Technol.*, vol. 6, no. 3, pp. 153–163, Sep 2024, doi: 10.3948/IJMT.2024/03020100.

[171] V. S. Chitale, H. Shashikumar, and D. M. Shrikhand Patil, "Heart Disease Detection Using Ensemble Machine Learning Technique on Various Risk Factors," *2nd IEEE Int'l Conf. Networks, Mobile Inf. Technol. NMICNT-2024*, pp. 1–8, Aug 2024, doi: 10.1109/NMICNT50024.2024.980891.

[172] I. D. Moustafa and R. J. Gray, "Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction," *Interventions*, vol. 15, no. 7, pp. 394–394, Jul 2024, doi: 10.3390/INT15070394.

[173] S. Zhou, "Generalization Error in Heart Disease Prediction: Comparative Analysis of MLP, Random Forest, and XGBoost on Chinese Heart Disease," *Sci. Technol. Engg. Environ. Prot.*, vol. 1, no. 1, Mar 2024, doi: 10.31137/STEEP-2024-010342.

[174] N. C. Harish, S. S. V. Lakshmi, M. L. Mughan, and K. V. Sure, "Cardiovascular Disease Prediction Using Ensemble Learning," *J. Engg. Comput. Technol.*, vol. 10, no. 1, pp. 1–10, Feb 2024, doi: 10.1109/CECEN50074.2024.1051007.

[175] G. B. Praveen, J. Prashanth, J. Kanthas, and S. Kumar, "Cardiovascular Disease Prediction Using Extreme Gradient Boosting Algorithm," *2023 2nd Int'l Conf. Adv. Comput. Infell. Comput. ICAC-2023*, pp. 1–6, Dec 2023, doi: 10.1109/ICAC9345.2023.10045219.

[176] D. Bhushan, V. Singh, R. Kaur, and S. Kumar, "Advancing Cardiovascular Disease Risk Assessment: A Comprehensive Study on Heart Attack Prediction Using XGBoost Models and Cross-Validation," *2024 2nd Int'l Conf. Adv. Comput. Sys. Eng. Technol. ICASCST-2024*, pp. 1–5, Sep 2024, doi: 10.1109/ICASCST50024.2024.980454.

[177] Y. Zhou, "Prediction and Analysis of Coronary Disease Diseases Based on MLBoost," *2024 10th IEEE 2nd Int'l Conf. Comput. Engg. Technol. ICCE-2024*, pp. 1877–1881, Jun 2024, doi: 10.1109/ICCE-2024.93918.2024.10018781.

[178] K. Aseeri and K. A. Ely, "Improving Heart Disease Prediction: A Comparative Study of Hypothesis Testing Techniques with Floating Algorithms," *2023 Int'l Conf. Adv. Inf. Syst. ICADS-2023*, pp. 1–6, Dec 2023, doi: 10.1109/ICADS5001.2023.9804524.

[179] S. P. Praveen, A. Venkatesh, C. S. Shashikumar, R. Prabhakar, and K. S. Venkateswaran, "Heart Disease Prediction using Neural Network: A Study of Ensemble and Convex Distance Metrics," *Proc. Int'l Conf. Smart. Civ. Engg. Constr. Infell. Comput. SCICON-SCION-2023*, pp. 236–240, Dec 2023, doi: 10.1109/SCICON-SCION5009.2023.9804520.

[180] S. N. Anil, A. M. S. Reza, S. Faraj, and A. R. Joneidi, "Klaarheid: Perceptrons Serangan Jaringan Menggunakan Model Machine Learning Klasifikasi Klasifikasi dan Klasifikasi Support Vector Mesin (SVM) pada Data Pendataan Kesehatan Nasional," *2024 1st Int'l Conf. Inf. Engg. Comput. Infell. Technol. ICIEC-2024*, pp. 1–6, Oct 2024, doi: 10.1109/ICIEC5001.2024.9804524.

[181] M. Z. Islam, M. A. Sharmin, and A. B. M. Karim, "Future value-informed prediction and selection process for predicting heart attack possibility and the level of severity by using machine learning models," *Int. J. Engg. Comput. Sci.*, vol. 13, pp. 107–115, 2023–2029, Jun 2024, doi: 10.18557/IJECSS.2017.0817-0813.

[182] P. A. Sharmin, S. Padmapriya, K. Sarker, L. Einan, and A. K. Tharun, "Heart Disease Detection Using Classification and Prediction," *2024 1st Int'l Conf. Comput. Environ. Measur. Techniques. ICCEMT-2024*, Apr 2024, doi: 10.1109/ICCEMT5004.2024.9804524.

[183] H. Liu, Y. Tian, and W. Yu, "Prediction of cardiovascular and cerebrovascular diseases based on machine learning models," *Appl. Comput. Engg.*, vol. 46, no. 1, pp. 35–44, Mar 2024, doi: 10.3124/95-9966-2024-02-0006.

[184] Nishad Arjun et al., "Improving Cardiovascular Disease Prediction through Comparative Analysis of Machine Learning Models," *J. Comput. Sci. Technol.*, vol. 6, no. 2, pp. 62–70, Apr 2024, doi: 10.3939/JCST.2024.020062.

[185] Y. Duan, L. Wu, Meng, and Y. Zhang, "Comparative analysis of supervised learning algorithms for prediction of cardiac related diseases," *Technol. Inf. Engg. Sci.*, vol. 26, no. 2, pp. 241–245, Apr 2024, doi: 10.3124/TIES-2024-020062.

[186] H. Elzain, M. S. M. Sadiq, and S. S. Sadiq, "Effect of feature (order) scaling, Optimal scaling, and SMOTE on the Performance of Cardiovascular Disease Prediction Models," *J. Mat. Sci. Engg. Comput.*, vol. 21, no. 1, pp. 88–102, Sep 2024, doi: 10.2929/V2111.2024.10552.

[187] M. Chandrashekar, R. Kiran, P. Vaishali, and S. Kruck, "Evaluation of Machine Learning Models for Cardio-Stroke Risk Assessment," *Int. J. Comput. Infell. Technol.*, vol. 1, pp. 95–98, Aug 2024, doi: 10.10019/ICIT5001.2024.10049637.

[188] M. N. Salihia, N. Ramadhaniyah, and E. S. Swastika, "Cardiovascular Wellness Through Technology: A Closer Look at Machine Learning Techniques," *Seadec Int'l Conf. Adv. Inf. Technol. ICAT-2024*, pp. 1–8, Jul 2024, doi: 10.1109/ICAT5003.2024.9804522.

[189] "Comparative Study of Machine Learning Algorithm in Detecting Cardiovascular Diseases," *Dayan.*, May 2024, doi: 10.2929/95-9965-2024-02-0005.

[190] R. G. et al., "Review on Heart Disease Prediction using Machine Learning Approaches," *Int. J. MobiComp. Res.*, vol. 6, no. 5, Oct 2024, doi: 10.3948/IJMR.2024.08050245.

[191] N. Ghaniyan et al. Ramadhan, Adiwijaya, W. Muliarni, and M. Ahsan Gunia, "Prediction of Cardio-Stroke Disease (CVS) in UPM Using Tree-based Ensemble Model," *2024 1st Int'l Conf. Comput. Engg. Technol. ICCE-2024*, pp. 1–6, Oct 2024, doi: 10.1109/ICCE5001.2024.9804527.

[192] K. Nihalani, B. U. Sehgal, B. F. Ahmad, M. Arjan, J. Ravi, and V. G. Krishnan, "Cardiovascular Disease Prediction Through Machine Learning Techniques," *Proc. 3rd Int'l Conf. Adv. Inf. Technol. ICAT-2024*, pp. 1–6, Jul 2024, doi: 10.1109/ICAT5003.2024.9804521.

[193] C. Yuan, X. Song, L. Liu, G. Wei, and J. Wang, "Machine Learning Model for Cardiovascular Disease Prediction," *Appl. Comput. Engg.*, vol. 2024, doi: 10.3124/95-9966-2024-02-0002.

[194] D. Bhushan, V. Singh, R. Kaur, and L. Sharma, "Advancing Cardiovascular Health Risk Assessment: A Comparative Study of Machine Learning Prediction Using XGBoost Models and Cross-validation," *2024 2nd Int'l Conf. Adv. Comput. Sys. Eng. Technol. ICASCST-2024*, pp. 1–5, Sep 2024, doi: 10.1109/ICASCST50024.2024.980454.

[195] T. Dong et al., "Developing Cardiovascular Risk Prediction Model for an Advanced Network Model with High-Level Random Effects," *Biostat.*, pp. 1–10, Oct 2024, doi: 10.3390/BIOSTAT5001.2024.100497.

[196] D. S. Land et al., "Effective cardiac disease classification using F-XGB and GWO approach," *Med. Eng. Phys.*, vol. 152, pp. 10429–10429, Sep 2024, doi: 10.1016/j.medengphy.2024.10429.

[197] N. Mostafaion et al., "Predicting risks for heart attack using machine learning methods," *Int. J. Pattern. Recogn. Artif. Intell.*, vol. 14, no. 6, pp. 6734–6734, Oct 2024, doi: 10.1191/14698093.2024.10047402.

[198] C. Yuan, X. Song, L. Liu, G. Wei, and J. Wang, "Machine Learning Model for Cardiovascular Disease Prediction," *Appl. Comput. Engg.*, vol. 2024, doi: 10.3124/95-9966-2024-02-0002.

[199] D. Bhushan, V. Singh, R. Kaur, and L. Sharma, "Advancing Cardiovascular Health Risk Assessment: A Comparative Study of Machine Learning Prediction Using XGBoost Models and Cross-validation," *2024 2nd Int'l Conf. Adv. Comput. Sys. Eng. Technol. ICASCST-2024*, pp. 1–5, Sep 2024, doi: 10.1109/ICASCST50024.2024.980454.

[200] T. Dong et al., "Developing Cardiovascular Risk Prediction Model for an Advanced Network Model with High-Level Random Effects," *Biostat.*, pp. 1–10, Oct 2024, doi: 10.3390/BIOSTAT5001.2024.100497.

[201] D. S. Land et al., "Effective cardiac disease classification using F-XGB and GWO approach," *Med. Eng. Phys.*, vol. 152, pp. 10429–10429, Sep 2024, doi: 10.1016/j.medengphy.2024.10429.

[202] N. Mostafaion et al., "Predicting risks for heart attack using machine learning methods," *Int. J. Pattern. Recogn. Artif. Intell.*, vol. 14, no. 6, pp. 6734–6734, Oct 2024, doi: 10.1191/14698093.2024.10047402.

[203] [203] "Implementing Stack with n-n Validation for Heart Disease Prediction," *Int. J. Pharm. Qual. Assur.*, vol. 15, no. 3, pp. 149–149, Sep 2024, doi: 10.2929/15.3.0001.

[204] R. K. P. Tepthai et al., "Stack with n-n Validation for Heart Disease Prediction," *Int. J. Pharm. Qual. Assur.*, vol. 15, no. 3, pp. 149–149, Sep 2024, doi: 10.2929/15.3.0001.

[205] A. H. Elmehri, A. Alabdullah, and B. A. Alameer, "A machine learning approach to cardiovascular disease prediction with ensemble feature selection," *Industr. Engg. Comput. Sci.*, vol. 33, no. 2, pp. 1030–1041, Feb 2024, doi: 10.1101/HEICS2023-100401.

[206] C. S. Naseem, A. R. Akbar, and C. S. Patrice, "Explainable AI Modeling in Predicting the Probability of Coronary Heart Disease," *Int'l J. Health. Care. Technol.*, vol. 16, pp. 978–982, Aug 2024, doi: 10.3233/SI-2024-0743.

- [202] "Mitigating Learning Bias in Healthcare Datasets," *KonfMaVoz*, Jul. 2024, doi: 10.31979/ETD_98WG-1H7W3.
- [203] A. Ogurcová, F. Švec, S. Basara, A. M. Alberán, and S. N. Queen, "Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases," *Diagnóstico*, vol. 14, no. 2, Jan. 2024, doi: 10.3390/DIAGNOSTIC1401014.
- [204] S. Nagarkarayya, G. Alankar, S. Shashidhar, and M. Sudhakar, "HMLP-CID-SIBM: A Hybrid Machine Learning Framework for Cardiac-related Disease Diagnosis Prediction Using the SMOTE Stacking Method," *Int. J. Artif. Syst. Network Syst.*, vol. 137, LNCS, pp. 571–581, Jan. 2023, doi: 10.1007/978-981-19-3010-4_47.
- [205] K. V. Somra, G. Pappagariyappa, and C. Tjortis, "Strategic Machine Learning Optimization for Cardiovascular Disease Prediction and High-Risk Patient Identification," *Algorithms*, vol. 17, no. 5, Apr. 2024, doi: 10.3390/ALGORITHMS1705018.
- [206] P. C. Karthik, S. Vijayakumar, and T. S. Pathanah, "Machine Learning Model Enabled with Data Optimization for Prediction of Coronary Heart Disease," pp. 1–5, Mar. 2024, doi: 10.1109/TQCB515942.2024.101049194.
- [207] A. Kachhava and J. Hitt, "An Intelligent System for Early Predictions of Cardiovascular Disease using Machine Learning," *J. Statist. Res.*, vol. 1, no. 3, Mar. 2023, doi: 10.4701/JJSRH_V1I3_269.
- [208] M. A. Safian et al., "Enhancing Clinical Validation for Early Cardiovascular Disease Prediction through Simulation, AI and Web Technology," *Diagnóstico*, vol. 14, no. 12, pp. 1308–1308, Jun. 2024, doi: 10.3390/DIAGNOSTIC1412100.
- [209] V. O. Igwevwe, F. I. Amadike, and E. Nwulu, "Cardiovascular Disease Risk Prediction for People Living With HIV Using Ensemble Deep Neural Network," *Int. Conf. Sci. Eng. Bus. Dev. Sustain. Dev. Goals*, LNBB/EGC, 2024, pp. 1–6, Apr. 2024, doi: 10.1109/SEB48160051.2024.10209052.
- [210] F. Li, J. Zhang, Y. Wang, H. Guo, W. Chen, and J. F. Paton, "Abstract 15589: Evaluating Methods to Mitigate Bias in Machine Learning-based Predictive Risk Models," *Circulation*, vol. 146, no. Suppl. 1, Mar. 2022, doi: 10.1161/CIRCULATIONAHA.121.042092.
- [211] C. J. Ellis et al., "Informed Cardiovascular Disease Prediction Modelling using Machine Learning Techniques: A Focus on Cardiomyopathy," *New Comput. Novel Syst.*, pp. 1–8, Apr. 2024, doi: 10.1088/0959-898X/2024/24/024541.
- [212] M. S. Alwan, M. A. Tolokter, and M. Z. Mamat, "Advancements in Cardiovascular Disease Detection Leveraging Data Mining and Machine Learning," *NanoRev*, Mar. 2024, doi: 10.1101/20240109.88422.
- [213] A. Petrucci, "Cardiovascular disease prediction with machine learning techniques," *J. Cardiol. Circ. Rev.*, vol. 17, no. 2, pp. 41–41, Apr. 2024, doi: 10.15406/JCCR.2024.17.00003.
- [214] T. Shabani and S. patil, "CVDFP: A Hybrid Feature Selection Method with Data-driven Approach for Cardiovascular Disease Prediction Framework using Machine Learning," Dec. 2022, doi: 10.21393/3083.3-85-2023.170.V1.
- [215] A. G. Tumasinge, O. S. Sydya, L. Magana, and T. J. Oyaro, "A weather feature dataset for prediction of short-term rainfall quantiles in Uganda," *Data Br.*, vol. 90, Sep. 2023, doi: 10.1016/j.bb.2023.109613.
- [216] "Climate Weather Dataset for Uganda," galvera, Jan. 2023, doi: 10.7910/DVN/POLYHP.
- [217] Alevconina Arapinda, Chris Vilches, lenora Andropoulos, and lenora Dimitsalopoulou, "Health promotion strategies for cardiovascular diseases in the community of Uganda: Policies, challenges, and potential interventions," *J. Clin. Res. Arch.*, vol. 10, no. 1, pp. 360–365, Sep. 2023, doi: 10.3074/4SRRKA.2023.10.1.00003.
- [218] M. Amrit, A. S. Narayanan, T. P. Dhingra Prasad, and R. I. Ani-Vignesh, "Cardiovascular Disease Prediction Using Bootstrap Algorithms," *Proc. Int. Conf. Cyber Power Comput. Technol. ICCPCT* 2023, pp. 20–25, Aug. 2023, doi: 10.1109/ICCPCT5811.2023.1040497.
- [219] M. E. Shirley, N. H. Kang, and G. Mori, "Step-by-Additive Explanation (SBAP) for Cardiovascular Disease Prediction," *Indust. Eng. Comput. Sci. Spec. Issue ICCS3.70M...-Proc.*, pp. 1429–1437, Jul. 2024, doi: 10.1109/SECS360600.2024.1042927.
- [220] M. Mukundanath and M. G. T. P. Kumar, "Interpretable Artificial Intelligence in Cardiovascular Health: An In-depth Analysis of Heart Disease Data," *Indians. J. Eng. Manag.*, vol. 08, no. 02, pp. 1–13, Feb. 2024, doi: 10.3984/IJEM2845.
- [221] K. Y. Wang, B. S. Kim, C. Lee, J. Park, and D. J. Chung, "CV-level prediction processor using DNA computing," *IEICE Electron. Express*, vol. 16, no. 9, pp. 529–534, Jan. 2009, doi: 10.1587/ELEX.6.529.
- [222] P. Singh, A. Singh, S. Lal, and V. Nak, "CVMDmagic: amenable AI based study for CVD risk detection in rural India," *ACM Int. Conf. Proceeding Ser.*, pp. 359–366, Mar. 2012, doi: 10.1145/2106673.2106716.
- [223] K. Gurumurthy, V. D. Anitha, Kumar, and K. Jayadevan, "An efficient cardio vascular disease prediction using multi-scale weighted feature fusion-based convolutional neural network with residual gains," *Comput. Commun. Comput. Res. Review (CCCR)*, vol. 27, no. 3, pp. 1–25, Apr. 2024, doi: 10.1080/1055688X.2024.2394425.
- [224] L. Foster, C. Katsoulis, and S. Tucker, "The anatomy of the great findings: a historical analysis of organizational," *Int. J. High Perform. Comput. Appl.*, vol. 15, no. 3, pp. 200–222, Sep. 2001, doi: 10.1177/104453200101500302.

AUTHORS



MR. SSIMBWABA BASHIR: Ssimbwa Bashir received his Bachelor's degree in Computer Science from Kampala International University in November, 2024. He is currently pursuing Data

Science at World Quant University and working on projects in machine learning and data analytics. His research interests include Data Engineering, Machine Learning, Engineering and Data Science & Analytics.



MR. WASIKE AZIZI ; Mr. wasike Azizi is a computer scientist, researcher, and lecturer with extensive expertise in Natural Language Processing, Machine Translation, Automata Theory, and Rule-Based Machine Translation. He is currently a PhD candidate in Computer Science (2022–2025) at Kampala International University, where he is conducting research in advanced computational linguistics and AI-driven language technologies. He holds a Master of Science in Computer Science (2018–2020) and a Bachelor of Science in Computer Science (2009–2012) from the same institution. Additionally, he earned a Postgraduate Diploma in Education (2021–2022), equipping him with skills in teaching ICT and Physics. With a strong background in academic and research, Mr. Wasike Azizi is dedicated to advancing AI-driven innovations in language processing and computational modelling while contributing to education and technology development.



Mr. Ssebadduka Jamir is a computer scientist, researcher, and lecturer specializing in advanced computing and artificial intelligence. He is currently a PhD candidate in Computer Science (expected completion: November 2025) at Kampala International University, where he is conducting research in emerging technologies and computational methodologies.

He holds a Master's Degree in Computer Science, further strengthening his expertise in data-driven systems and software development. In addition to his academic pursuits, Mr. Ssebadduka Jamir is a lecturer at the International University of East Africa (IUEA) and the Islamic University in Uganda (IUU), where he is committed to shaping the next generation of computer scientists through teaching and research.

His work focuses on innovative AI applications, system optimization, and technological advancements, contributing significantly to both academia and industry.

PRIMARY SOURCES

- | | | |
|---|--|--------|
| 1 | ijrpr.com
Internet Source | $<1\%$ |
| 2 | "Proceedings of the Sixth International Conference on Trends in Computational and Cognitive Engineering", Springer Science and Business Media LLC, 2026
Publication | $<1\%$ |
| 3 | www.irjmets.com
Internet Source | $<1\%$ |
| 4 | carcinogenesis.com
Internet Source | $<1\%$ |
| 5 | www.isteonline.in
Internet Source | $<1\%$ |
| 6 | www.ijraset.com
Internet Source | $<1\%$ |
| 7 | www.drngpasc.ac.in
Internet Source | $<1\%$ |
| 8 | pubmed.ncbi.nlm.nih.gov
Internet Source | $<1\%$ |
-

9	pubs.thesciencein.org Internet Source	<1 %
10	nusantarahasanajournal.com Internet Source	<1 %
11	Submitted to Sbec High School Student Paper	<1 %
12	The Electronic Library, Volume 27, Issue 6 (2009-12-15) Publication	<1 %
13	www.geeksforgeeks.org Internet Source	<1 %
14	www.techscience.com Internet Source	<1 %
15	Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security - Volume 2", CRC Press, 2023 Publication	<1 %
16	Submitted to Rajiv Gandhi Institute of Technology, Kottayam Student Paper	<1 %
17	www.jpost.com Internet Source	<1 %
18	www.mdpi.com Internet Source	<1 %

- 19 Beat Knechtle, Sasa Duric, Volker Scheer, Pantelis T. Nikolaidis et al. "Resilience against exercise-related coronary atherosclerosis: A case study in a master athlete participating in 500 marathons", Sports Medicine and Health Science, 2025 <1 %
- Publication
-
- 20 www.nature.com <1 %
- Internet Source
-
- 21 S. Vijayalakshmi, Alwin Joseph, Naived George Eapen, Balamurugan Balusamy, Jagjit Singh Dhatterwal, Kuldeep Singh Kaswan. "Smart Systems for Disease Prediction - Advancements, Applications and Challenges", CRC Press, 2025 <1 %
- Publication
-
- 22 continentalhospitals.com <1 %
- Internet Source
-
- 23 ojs.ijemd.com <1 %
- Internet Source
-
- 24 M Nivedhitha, Durai Raj Vincent. "Deep Belief Network-based Approach for Detecting Alzheimer's disease using the Multi-omics Data", Computational and Structural Biotechnology Journal, 2023 <1 %
- Publication
-

- 25 Md. Alamin Talukder, Amira Samy Talaat, Mohsin Kazi. "HXAI-ML: A hybrid explainable artificial intelligence based machine learning model for cardiovascular heart disease detection", Results in Engineering, 2025
Publication <1 %
- 26 www.medrxiv.org Internet Source <1 %
- 27 "Bioinformatics and Biomedical Engineering", Springer Science and Business Media LLC, 2026
Publication <1 %
- 28 Mohammad Amin Kuhail, Mostafa Mohamad, Rawad Hammad, Mohammed Bahja. "Unleashing User Innovation - Multidisciplinary Perspectives on End-User Development and Generative AI", CRC Press, 2025
Publication <1 %
- 29 Selorm Adablanu, Utpal Barman, Dulumani Das. "Advancing deep learning for automated stroke detection: a review", Brain Hemorrhages, 2025
Publication <1 %
- 30 adamdjellouli.com Internet Source <1 %

- 31 jogh.org Internet Source <1 %
- 32 store.samhsa.gov Internet Source <1 %
- 33 "Artificial Intelligence and System Engineering", Springer Science and Business Media LLC, 2025 Publication <1 %
- 34 Arun PV, Jocelyn Chanussot, B Krishna Mohan, D Nagesh Kumar, Alok Porwal. "Explainable AI for Earth Observation Data Analysis - Applications, Opportunities, and Challenges", CRC Press, 2025 Publication <1 %
- 35 Ibomoiye Domor Mienye, George Obaido, Nobert Jere, Ebikella Mienye et al. "A survey of explainable artificial intelligence in healthcare: Concepts, applications, and challenges", Informatics in Medicine Unlocked, 2024 Publication <1 %
- 36 R. Subathra, V. Sumathy. "A smart CardioSenseNet framework with advanced data processing models for precise heart disease detection", Computers in Biology and Medicine, 2025 Publication <1 %

- 37 Samet Aymaz. "Boosting medical diagnostics with a novel gradient-based sample selection method", Computers in Biology and Medicine, 2024 <1 %
- Publication
-
- 38 Vibhav Sachan, Shahid Malik, Ruchita Gautam, Kumar Parvin. "Advances in AI for Biomedical Instrumentation, Electronics and Computing", CRC Press, 2024 <1 %
- Publication
-
- 39 Vincent Zibi Mohale, Ibidun Christiana Obagbuwa. "A systematic review on the integration of explainable artificial intelligence in intrusion detection systems to enhancing transparency and interpretability in cybersecurity", Frontiers in Artificial Intelligence, 2025 <1 %
- Publication
-
- 40 blueholetek.blogspot.com <1 %
- Internet Source
-
- 41 docplayer.net <1 %
- Internet Source
-
- 42 fastercapital.com <1 %
- Internet Source
-
- 43 siit.ugatu.su <1 %
- Internet Source

- 44 www.frontiersin.org <1 %
Internet Source
-
- 45 www.hasanbalik.com <1 %
Internet Source
-
- 46 "Advanced Computing and Intelligent Technologies", Springer Science and Business Media LLC, 2026 <1 %
Publication
-
- 47 Akeem Temitope Otapo, Alice Othmani, Ghazaleh Khodabandelou, Zuheng Ming. "Prediction and detection of terminal diseases using Internet of Medical Things: A review", Computers in Biology and Medicine, 2025 <1 %
Publication
-
- 48 Ameh, Emmanuel Esem. "Multimodal Deep Learning Algorithms for Predictive Modeling of Cardiovascular Diseases Onset and Progression", Capitol Technology University, 2025 <1 %
Publication
-
- 49 Anjan Bandyopadhyay, Tanvir Habib Sardar, Saurav Mallik, Ruhul Amin Hazarika, Mahendra Kumar Gourisaria. "AI and Data Engineering for Healthcare - Real-World Applications and Case Studies", CRC Press, 2025 <1 %
Publication
-

- 50 Lin Huang, Mohammad Reza Rajati. "Rule-Extraction from Soft Decision Trees", 2018 International Conference on Big Data and Artificial Intelligence (BDAI), 2018
Publication <1 %
- 51 Ravichander Janapati, Usha Desai, Steven Fernandes, Rakesh Sengupta, Shubham Tayal. "Applied Artificial Intelligence and Machine Learning Techniques for Engineering Applications", CRC Press, 2025
Publication <1 %
- 52 Vivek Kumar Singh, Jitendra Kumar, Deepika Saxena, Abhishek Verma, T. Akilan, Ashutosh Kumar Singh. "Machine Vision Analysis in Industry 5.0 - Fundamentals, Applications, and Challenges", CRC Press, 2025
Publication <1 %
- 53 datamites.com <1 %
Internet Source
- 54 medcraveonline.com <1 %
Internet Source
- 55 scity.org <1 %
Internet Source
- 56 www.journalimcms.org <1 %
Internet Source

- 57 Adedayo E. Ojo, Dike B. Ojji, Diederick E. Grobbee, Mark D. Huffman, Sanne A. E. Peters. "The Burden of Cardiovascular Disease Attributable to Hypertension in Nigeria: A Modelling Study Using Summary-Level Data", Global Heart, 2024
Publication <1 %
-
- 58 Maithri Bairy, Krishnaraj Chadaga, Niranjana Sampathila, R. Vijaya Arjunan, G. Muralidhar Bairy. "An Explainable Analytical Approach to Heart Attack Detection Using Biomarkers and Nature-Inspired Algorithms", Healthcare Analytics, 2025
Publication <1 %
-
- 59 Suresh S, Dhanalakshmi S. "Tuberculosis prediction: performance analysis of machine learning models for early diagnosis and screening using symptom severity level data", International Journal of Basic and Applied Sciences, 2025
Publication <1 %
-
- 60 "Artificial Intelligence in Human Health and Diseases", Springer Science and Business Media LLC, 2025
Publication <1 %
-
- 61 "Artificial Intelligence: Theory and Applications", Springer Science and Business Publication <1 %

Media LLC, 2025

Publication

62

"Graph Neural Networks for Neurological Disorders", Springer Science and Business Media LLC, 2025

<1 %

Publication

63

Arvind Dagur, Sohit Agarwal, Dhirendra Kumar Shukla, Shabir Ali, Sandhya Sharma. "Artificial Intelligence and Sustainable Innovation - Volume 1", CRC Press, 2026

<1 %

Publication

64

dos Santos, Ricardo Bruno Barbeiro. "Improving the Robustness of Multimodal AI with Asynchronous and Missing Inputs", Universidade NOVA de Lisboa (Portugal)

<1 %

Publication

Exclude quotes

On

Exclude matches

Off

Exclude bibliography

On