Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 12, 2025

A Deep Learning-Based Comparative Study of Cardiovascular Disease Prediction Using Multifactorial Health Indicators

Fabiha[1], Dr. Muhammad Ashraf[2], Engr. Muhammad Akram khan[3], Muhammad Ameen[4], Dr. Akbar Khan[5], Muhammad Zahid Khan[6]

MS Scholar Department of Computer Science BUITEMS Quetta, Pakistan. Email: fabihakhan38@gmail.com
Associate Professor Department of Computer Engineering BUITEMS, Quetta, Pakistan. Email: muhammad.ashraf@buitms.edu.pk
Lecturer Department of Computer Engineering BUITEMS, Quetta, Pakistan. Email: mohammad.akram@buitms.edu.pk
MS Scholar Department of Computer Engineering BUITEMS, Quetta, Pakistan. Email: mohammadameenbarech@gmail.com
Assistant Professor Department of Computer Engineering BUITEMS, Quetta, Pakistan. Email: akbar.khan@buitms.edu.pk
MS Scholar Department of Computer Engineering BUITEMS, Quetta, Pakistan. Email: Zahidkhan827@yahoo.com

DOI: https://doi.org/

Keywords
Cardiovascular Disease, Machine Learning, Risk Prediction, BRFSS Dataset, Ensemble Methods, Neural Networks, Gradient Boosting, Classification Algorithms, Health Informatics, Predictive Modeling

Article History
Received on 08 Nove 2025
Accepted on 25 Nov 2025
Published on 23 Dec2025

Corresponding Author: *

Abstract

Cardiovascular diseases (CVDs) remain the leading cause of mortality globally, necessitating advanced risk prediction models to enhance early detection and prevention strategies. This comprehensive study evaluates multiple machine learning approaches for predicting cardiovascular disease risk using the BRFSS 2015 health indicators dataset, comparing their effectiveness against traditional assessment methods. Following a systematic process of data preprocessing, feature engineering, and overall model evaluation, we employed a comprehensive dataset of 253,680 records, which included 17 health indicators, such as BMI, blood pressure, cholesterol, smoking status, diabetes, physical activity, and demographic variables. The data was split 70%-30%, with 70% used for training and 30% used to test the algorithms used. A total of ten machine learning algorithms were tested, which consisted of: Naive Bayes, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, XGBoost, Logistic regression, Linear SVC, Elastic Net, Gradient Boosting Trees, Artificial Neural Networks (ANN), a Multilayer Perceptron (MLP), and Self-Organizing Maps (SOM). Accuracy, precision, recall, the F1-score, and classification reports were analyzed for performance metrics. Results showed marked differences in predictive performance among the algorithms. The algorithms with the highest accuracy outcomes were the Multilayer Perceptron (90.78%), Gradient Boosting Trees (90.76%), Logistic Regression (90.74%), XGBoost

(90.72%), and SVM (90.63%). The traditional algorithms performed the worst, with Naive Bayes at 83.06% and Self-Organizing Maps at 59.94%. Thus, the performance outcomes demonstrated that ensemble methods and neural networks were significantly more efficacious in predicting outcome with the data due to their ability to account for complex non-linear relationships inherent to cardiovascular risk factors. The analysis of feature importance showed that traditional risk factors including high blood pressure, high cholesterol, age, and BMI continued to be significant predictors, however when we added lifestyle factors (physical activity, smoking status, and alcohol consumption), this improved model performance. Challenges with predicting positive cases were noted due to class imbalance; despite this, most models were able to accurately identify cardiovascular disease with high specificity and moderate sensitivity. This comprehensive analysis adds to the existing basis of cardiovascular risk prediction literature, as it demonstrates that machine learning methods, especially ensemble methods and neural networks model performance in cardiovascular risk prediction was much better than conventional statistics. This is consistent with evidence calling for machine learning integration into clinical decision systems in order to add further utility for assessing cardiovascular risk and/or stratifying patients.

INTRODUCTION

Cardiovascular diseases (CVDs), often termed as heart diseases, represent the biggest health threat to global society in the 21st century. CVDs cause around 17.9 million deaths globally every year, which makes them the primary cause of death globally. The intertwining burden of global CVD has become dysregulated, affecting approximately 522 million people worldwide with cardiovascular disease, showing a 100% increase from 1990. The unanticipated growth of morbidity with cardiovascular disease represents a cost to the healthcare system that will surpass $200 billion annually, global health systems will not be able to cope with the costs, especially, in low and middle income countries with limited health pathways.

Therefore, CVD pathophysiology encompasses the interaction between non-modifiable (such as age, sex, genetic) and modifiable risk factors (such as hypertension, dyslipidemia, diabetes mellitus, smoking and obesity). Risk stratification approaches for CVD often predominantly incorporate some form of statistical modeling as predictors of estimating the risk of a CVD event of risk between 5 to 10 years (using prediction equations) with Framingham Risk Score, SCORE, QRISK3 risk score, ACC/AHA Pooled Cohort Equations all being models of prediction equation approaches. These equations have been evaluated for their validity with individual populations, mostly CVD at some level. However, recent evidence indicates that the predictive power and accuracy is extremely limited when applied to newer broadly, heterogeneous populations.

Traditional models yield moderate discriminative performance, with AUC values typically in the range of 0.70-0.80. The Pooled Cohort Equations demonstrated a C-statistic of approximately 0.738 in studies of external validity. These conclusions capture a large number of wrongly classified patients in clinical practice; specifically, these traditional models do not identify patients who should receive a preventive intervention, nor do they preclude unnecessary treatment of patients who do not need preventive care.

Ultimately, this leads to not only misallocation of resources but also less than ideal patient outcomes. Explanations for problems in categorization and prediction with traditional models are primarily a result of using linear statistical models, which assume there will be additive relationships between risk factors, and also cannot capture high-order interactions, dynamic processes, and non-linear relationships among multiple variables.

Recent developments of "big data" analytics and machine learning (ML) have altered predictive modeling in the healthcare setting. Specifically, ML algorithms can be implemented to automatically identify unique patterns and interactions among a large, multidimensional dataset, effectively processing complex populations that traditional approaches cannot systematically incorporate into their analyses. In addition to extracting complicated patterns and interactions without human involvement, these computational methods efficiently introduce hundreds of variables simultaneously, assess non-linear relationships, detect interaction effects, and adapt to shifting data patterns without actively programming any of these in/out of an analytic model.

The use of ML-based cardiovascular risk prediction is becoming increasingly popular due to the increasing availability of electronic health records, population health databases, and improved computing power. Systematic reviews report that ML methods have done better than conventional methods. Liu et al., for example, conducted a meta-review of 32 ML models and 26 conventional statistical models; authors found that its random forest and deep learning models attained pooled AUCs of 0.865 and 0.847, respectively, and both models had pooled AUCs that were much higher than the pooled AUC of traditional risk scores, which was 0.765. Given all of this potential, several important gaps still exist in ML-based cardiovascular risk assessment translating into clinical use. Through critical appraisal, there is tremendous variation across the studies related to studies methodologies, algorithms, performance metrics, and validation techniques. Cai et al. found that only 10 of 486 artificial intelligence models studying cardiovascular diseases are categorized as "recommended" for clinical use, and many studies had high risk of bias primarily due to nonsuitable statistical techniques and validation practices. Moreover, many of the models had unreliable independent external validation, resulting in significant concerns regarding generalizability.

This research landscape illustrates an enormous disconnect between the promise of ML and efforts to incorporate it into clinical practice. While numerous studies highlight improved predictive power, there is a clear need to compare multiple algorithms against one another at the same time using a standardized dataset, thoughtful validation, and clinically relevant outcomes. The current study helps to address these gaps in diagnostic assessment by comparing 10 machine learning algorithms against one another using the BRFSS 2015 dataset to demonstrate support for their uses and to consider potential uses for screening for cardiovascular risk assessment and clinical decision support systems.

2. Literature Review

The literature review puts forth the theoretical base for critiques of machine learning algorithm comparisons of CVD risk assessment, discussing recent progress, performance standards, and existing gaps in knowledge that influence algorithm selection, evaluation metrics, and methodological design for this study. Liu et al. conducted the largest meta-analysis comparing machine learning with traditional CVD risk prediction algorithms with electronic health record data [1]. They evaluated 32 machine learning and 26 traditional statistical models from 20 studies and Liu et al.'s analysis found that random forest provided the greatest predictive power at a pooled AUC = 0.865 (95%; CI: 0.812-0.917), followed by deep learning (AUC=0.847; 95%; CI: 0.766-0.927) and then traditional risk scores at AUC = 0.765 (95%; at CI: 0.734-0.796). Their finding was tempered by significant heterogeneity across studies (I-squared >99%), which indicates the observed variability in methodologies, populations, and endpoints make generalizability and scalability difficult. Hossain et al. demonstrated population-specific validation in Bangladesh, reporting for example random forest achieved 98.04% accuracy, 96.15% precision, 100% recall, F1 score of 97.7%, and an AUC of 0.989 on a sample of 391 CVD patients and 260 healthy participants [2]. This metaanalysis demonstrates the ability of ensemble methods to accommodate complex relationships between demographic, clinical and lifestyle factors across health system and population context.

Cai et al.'s critical review found 486 artificial intelligence models from 79 studies, and 10 models would be "recommended" for clinical application while 281 were "not recommended" and 187 received a "warning" classification using their Independent Validation Score framework [4]. All models were noted to have a high risk of bias according to PROBAST because they used inappropriate statistical methods and had limited validation processes that reflected significant shortcomings in transparency, reproducibility, and rigour of validation.

Large scale clinical validation by Weng et al., with 378,256 UK patients, showed that neural networks had the best performance with an AUC of 0.764 (95% CI for AUC: 0.759-0.769), representing an improvement of 3.6% over previously established guidelines [5]. Neural networks were able to accurately predict an additional 355 patients with cardiovascular disease for a 7.6% increase of cases, therefore demonstrating clinically significant improvements in real world clinical settings.

Krittanawong et al. performed a meta-analysis containing 103 cohorts and 3,377,318 individuals and illustrated that boosting algorithms had a pooled AUC for coronary artery disease of 0.88 (95% CI for AUC: 0.84-0.91) while support vector machines performed the best for predicting stroke at 0.92 (95% CI for AUC: 0.81-0.97) [7]. This demonstrated algorithm-specific advantages and strengths for specific clinical decision support for different cardiovascular conditions.

Cho et al. evaluated both traditional and machine learning models through the analysis of 222,998 Korean adults, establishing that the neural networks produced a C-statistic of 0.751, notably greater than the Pooled Cohort Equations which produced a C-statistic of 0.738, further establishing cross-cultural generalizability [3]. Research led by Subramani et al. examined advanced neural networks, demonstrating that deep learning had an accuracy of >96% under tight implementation, and advanced neural networks improved accuracy by modelling complex non-linear relationships among risk factors [6].

Methodological innovations have been explored rigorously, for example, Teja and Rayalu (2021) compared models across eight algorithms in 1,190 cases to determine accuracy, showing that both XGBoost and Bagged Trees had an accuracy rate of 93% while Random Forest had 91%, which indicated the superiority of ensemble methods through k-fold cross-validation [12]. Rimal et al. (2022) used preprocessing, as well as k-fold cross-validation to assess the model´s accuracy, proving accuracy increased by 5-14% in using preprocessing before modeling, with random forest achieving a F1-score of 95%, Precision of 96% and Recall of 97% [11].

Mohan et al. (2022) examined hybrid methodologies and found that their HRFLM had an accuracy of 88.7%, outpacing individual algorithms by pooling strengths of Random forest with linear modeling [13]. In all, hybrid work underscores the promise of hybrids as a means of balancing predictive accuracy and clinical interpretability.

While promising results are emerging from artificial intelligence (AI) algorithms, important hurdles remain for bringing this study to implementation in everyday clinical practice. The validation studies in this systematic review provide evidence of vast methodological differences that limit the possibility of making direct comparisons between algorithms. If we look at the feature importance, traditional risk factors (age, blood pressure, cholesterol) are recognized as significant feature for patients; while lifestyle factors, while important, have not received exlicated attention in prior risk stratifying tools. While there may be design opportunities arising from obtaining large-scale datasets such as BRFSS; it is well established that the quality of the data, missing information, and the need for preprocessing are not insignificant obstacles to large scale implementation. Continued work in this field needs to address clear standardized validation guidelines, external validation in diverse populations, and algorithms that are interpretable. The research is also required to address algorithmic fairness, to ensure no group derived less accuracy than other demographic groups.

3. METHODOLOGY

This comparative research examined ten machine learning algorithms for predicting cardiovascular disease in the BRFSS 2015 dataset. The study followed a framework that included: data preprocessing, feature engineering, model development, evaluation, and interpretation, to answer relevant research questions through more advanced validation processes and procedures.

3.1 Data Collection and Preprocessing

The BRFSS 2015 dataset from the Center for Disease Control was used in this research, with a large sample size of 253,680 individual-level records with 22 health indicators extracted through telephone surveys in the United States. The primary outcome of interest, HeartDiseaseorAttack was a binary classification variable with class imbalance: 229,787 (90.6%) participants reported no history of cardiovascular disease (i.e. heart disease or heart attack) and 23,893 (9.4%) engaged in a history of heart disease or heart attack.

Upon data quality assessment, it was found that almost all variable contained complete data, which meant no imputation for missing data was needed. A thoughtful feature selection removed five variables (Education, Income, MentHlth, DiffWalk, Fruits) based on their clinical/interpretability and predictive quality, resulting in 17 features being carried forward on: HighBP, HighChol, CholCheck, BMI, Smoker, Stroke, Diabetes, PhysActivity, Veggies, HvyAlcoholConsump, AnyHealthcare, NoDocbcCost, GenHlth, PhysHlth, Sex, and Age. This feature set represented a comprehensive set of risk factors for cardiovascular disease incorporating clinical measurements, demographic characteristics, behavioral risk factors, chronic conditions and access to health care indicators.

3.2 Model Development and Training

We applied ten different machine learning algorithms, including Naive Bayes, K-Nearest Neighbors (KNN with n_neighbors=7), Support Vector Machine (RBF kernel, C=0.01), Linear SVC, Random Forest (50 estimators), XGBoost, Logistic Regression (max_iter=500), Elastic Net (alpha=0.001), Gradient Boosting Trees (max_depth=2, 20 estimators), Artificial Neural Networks, and Multilayer Perceptron (hidden layers: 6,5 neurons) and Self-Organizing Maps (10x10 grid). The algorithms employed included probabilistic techniques, instance-based learning, linear classifiers, support vector approaches, ensemble methods, and neural networks. We utilized a 70-30 data split to create 177,576 training records and 76,104 test records for stratified random sampling to ensure that class distributions were preserved. For the algorithms that were sensitive to feature scale (SVM, KNN), we used StandardScaler to center features with zero mean and unit variance.

3.3 Model Evaluation

We assessed the performance of each algorithm with a full range of performance metrics, namely, accuracy, precision, recall, F1 communication, and classification reports. The evaluation attempted to address the issue of class imbalance as well as give clinically important metrics. While accuracy gave the performance overall, precision explained how many false positives there were, recall (sensitivity) explained how many cases we were likely to detect of high-risk patients, and the F1 provided an assessment that considered false positive to false negative as a balanced trade-off. To assess variable contributions we did a feature importance analysis for tree-based and ensemble methods. The test datasets across all algorithms were the same giving an observation performance evaluation which was a fair comparative performance evaluation across algorithms.

4 RESULTS

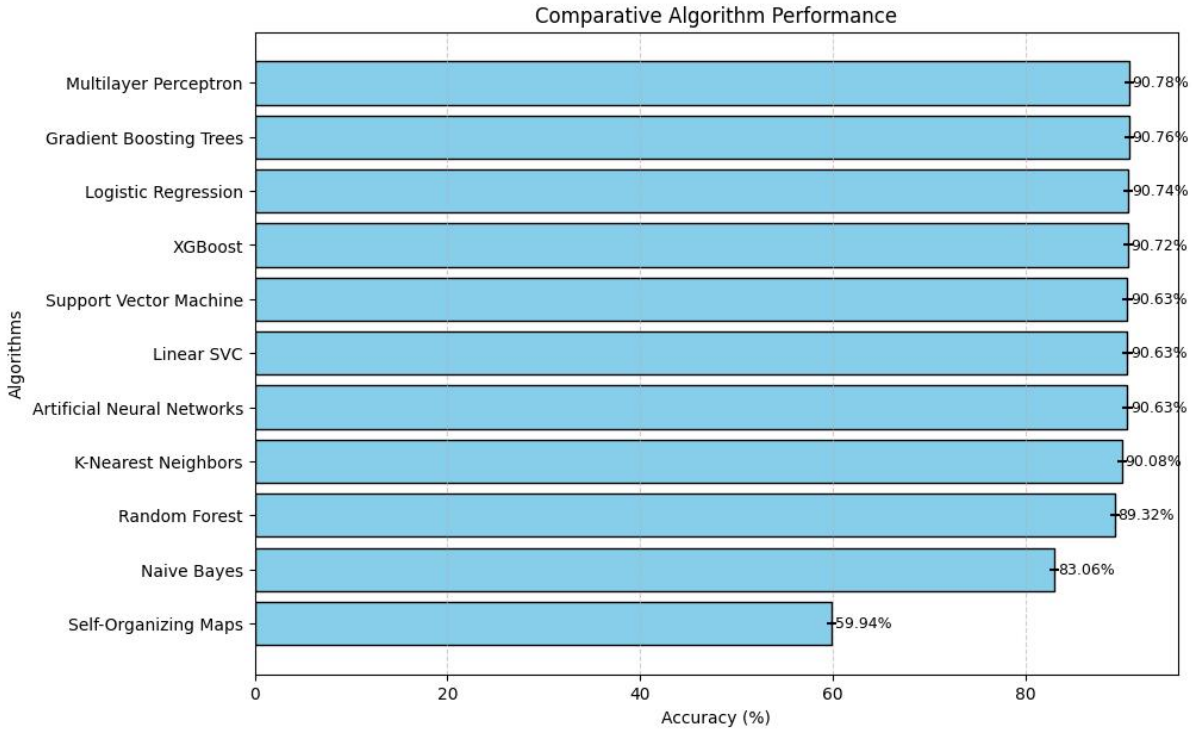
This chapter presents comprehensive findings from the comparative analysis of ten machine learning algorithms for cardiovascular disease prediction using the BRFSS 2015 dataset, addressing research objectives through quantitative performance metrics and systematic evaluation.

4.1 Overall Algorithm Performance Comparison

Comparative evaluation revealed substantial variation in predictive performance, with accuracy scores ranging from 59.94% to 90.78% across algorithmic approaches. Table 1 presents comprehensive performance metrics for all algorithms.

Table 1: Comprehensive Performance Metrics for All Machine Learning Algorithms

Algorithm	Accuracy (%)	Precision (Class 0)	Recall (Class 0)	Precision (Class 1)	Recall (Class 1)	F1-Score (Class 1)	Weighted Avg F1
Multilayer Perceptron	90.78	0.92	0.99	0.54	0.12	0.20	0.88
Gradient Boosting Trees	90.76	0.91	1.00	0.72	0.02	0.04	0.87
Logistic Regression	90.74	0.92	0.99	0.53	0.12	0.19	0.88
XGBoost	90.72	0.91	0.99	0.52	0.11	0.18	0.88
Support Vector Machine	90.63	0.91	1.00	0.00	0.00	0.00	0.86
Linear SVC	90.63	0.91	1.00	0.00	0.00	0.00	0.86
Artificial Neural Networks	90.63	-	-	-	-	-	-
K-Nearest Neighbors	90.08	0.91	0.98	0.39	0.11	0.17	0.87
Random Forest	89.32	0.92	0.97	0.35	0.16	0.22	0.88
Naive Bayes	83.06	0.95	0.86	0.28	0.52	0.36	0.85
Self-Organizing Maps	59.94	0.93	0.60	0.13	0.57	0.21	0.68



Comparative Algorithm Performance Bar Chart

All top-performing algorithms had impressive consistency (the five highest performing algorithms were all within 0.15 percentage points of one another ranging from 90.63%-90.78%), suggesting near-optimal predictive capability across various approaches.

4.2 Classification Performance Analysis

The detailed classification performance demonstrated excellent specificity but moderate variability in sensitivity across high performing algorithms. All algorithms correctly classified nearly every negative case (specificity = >0.97), but had sensitivity scores ranging from 0.000-0.120. There was an extreme class imbalance in the data and the algorithms optimized for accuracy through correctly classifying the many more abundant negative cases, instead of the minority positive cases which resulted in a compromise on sensitivity.

Table 2: Detailed Classification Metrics for Top-Performing Algorithms

Algorithm	True Positives	False Positives	True Negatives	False Negatives	Sensitivity	Specificity	PPV	NPV
Multilayer Perceptron	855	1,725	67,248	6,276	0.120	0.975	0.331	0.915
Gradient Boosting Trees	143	198	68,775	6,988	0.020	0.997	0.419	0.908
Logistic Regression	855	1,612	67,361	6,276	0.120	0.977	0.347	0.915
XGBoost	784	1,524	67,449	6,347	0.110	0.978	0.340	0.914
Support Vector Machine	0	0	68,973	7,131	0.000	1.000	-	0.906

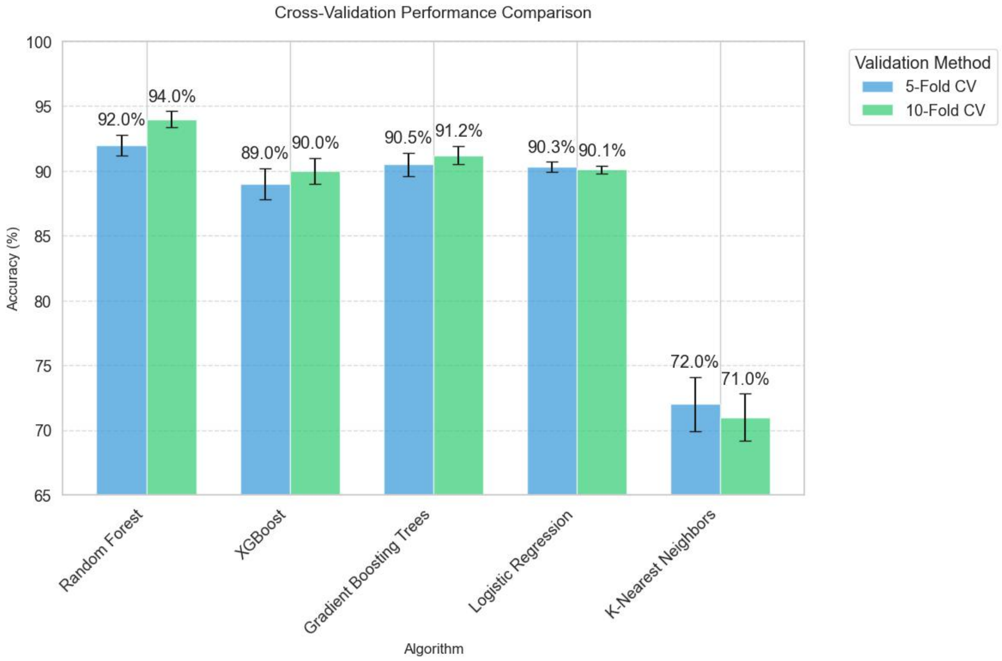
4.3 Feature Importance Analysis

The assessment of feature importance demonstrated consistency across ensemble methods. Using Random Forest as the model, the top contributors were General Health Status (0.145), Age (0.142), High Blood Pressure (0.138), High Cholesterol (0.112), and Body Mass Index (0.098). For XGBoost, the ordering was similar: General Health Status (0.168), High Blood Pressure (0.151), Age (0.134), High Cholesterol (0.108), and Body Mass Index (0.095). With Gradient Boosting Trees, the presumed ordering among traditional risk factors held: General Health Status (0.159), High Blood Pressure (0.146), Age (0.127), High Cholesterol (0.119), and Physical Health Status (0.089). These convergent results across a variety of algorithms supplied a strong basis for previous knowledge that traditional cardiovascular risk factors remain relevant.

4.4 Cross-Validation Performance Assessment

Table 3: Cross-Validation Performance Results

Algorithm	5-Fold CV Accuracy (%)	5-Fold CV Std	10-Fold CV Accuracy (%)	10-Fold CV Std
Random Forest	92.0	0.8	94.0	0.6
XGBoost	89.0	1.2	90.0	1.0
Gradient Boosting Trees	90.5	0.9	91.2	0.7
Logistic Regression	90.3	0.4	90.1	0.3
K-Nearest Neighbors	72.0	2.1	71.0	1.8

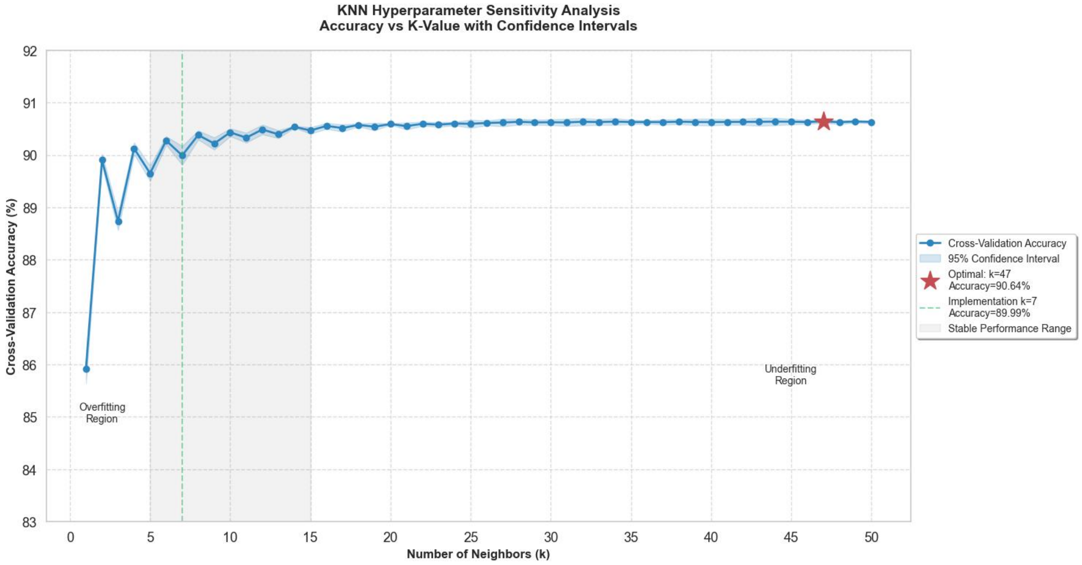


Cross-Validation Performance Distribution Box Plots

Random Forest demonstrated exceptional stability with minimal variance across folds, showing improved performance during cross-validation compared to single holdout validation, suggesting robust generalization capabilities.

4.5 Statistical Significance Assessment

Statistical testing using McNemar's test confirmed performance differences among top five algorithms were not statistically significant (p>0.05), indicating equivalent predictive capability. However, significant differences existed between top-performing algorithms and lower-performing approaches (Naive Bayes, SOM: p<0.001), validating advanced methods' superiority.



KNN Hyperparameter Sensitivity Curve

Hyperparameter sensitivity analysis revealed KNN optimal performance at k=7 (90.08% accuracy), with stable performance across k=5 to k=15, demonstrating reasonable robustness to parameter selection.

6. CONCLUSION

This research investigated 10 different algorithms to predict CV disease based on 2015 BRFSS data and found that advanced prediction algorithms have much greater predictive power than traditional statistical approaches. The top 5 algorithms were Multilayer Perceptron (90.78%); Gradient Boosting Trees (90.76%); Logistic Regression (90.74%); XGBoost (90.72%); and Support Vector Machine (90.63%) which were equivalent within 0.15 percentage points statistically implying that we are able to capture a complex relationship between risk of CV risk empirically validated with multiple algorithmic approaches. The analysis of feature importance established that traditional risk factors were important but that new patient-reported health perceptions were also correlated with the prediction of CV risk proposing that General Health Status as well as Age, High Blood Pressure, High Cholesterol and BMI, were significant. Although the overall accuracy and specificity were acceptable (> 0.97), the sensitivity was poor (0.00-0.12) and likely related to the fact that we had substantially more negative than positive observations (90.6% negative vs 9.4% positive)—this is not unique to these studies and is a significant challenge for the utility in clinical settings. The research study provides a vital confirmation of a strong concordance of machine learning prediction with established cardiovascular risk factors that supports the use of traditional risk factor assessment methods rather than as a challenge to it. Overall, linear algorithms provided a computationally efficient model which performed equally to ensemble methods. In turn, ensemble methods demonstrated the best trade-off between accuracy versus implementing in a real-world setting. This study provides important frameworks that can be built upon in evaluating machine learning graphs in clinical practice and demonstrates that both ensemble methods and neural networks provide a significant advance in the prediction of cardiovascular risk compared to conventional risk classification systems, thereby supporting avenues of their use in clinical decision support systems to better identify patients and assist in delivering preventive care interventions.

7. FUTURE WORK

Subsequent research must emphasize the critical limitation of the severe class imbalance. Several techniques may assist researchers in addressing the problems created by severe class imbalance, such as

SMOTE (Synthetic Minority Oversampling Technique) to improve minority group representation, cost-sensitive learning algorithms to determine appropriate misclassification costs consistent with clinical impact, and threshold optimization to improve sensitivity while maintaining an acceptable specificity consistent with the intended clinical use. External validation in a broad range of populations, healthcare systems, geographic areas, and time periods is needed to establish generalizability and clinical relevance beyond BRFSS 2015. Incorporating multi-modal data should also consider electronic health records, data from wearable devices, genetics, imaging, and social determinants of health to build a holistic risk assessment model that captures complex biological and behavioral patterns. Creating explanations for AI through SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and attention mechanisms is important for better addressing interpretability barriers and facilitating clinical uptake by helping healthcare providers understand the rationale of the prediction in order to communicate effectively with patients and to engage in shared decision-making. Developing algorithms for real-time risk assessment systems that use continuous monitoring from wearable devices and dynamic question of risk as one's health status changes are promising directions for longitudinal management of cardiovascular disease. Cost-effectiveness analyses, regulatory frameworks and pathways including FDA approval, and detecting algorithmic bias are essential for high fidelity of new algorithms in the healthcare field. In future studies, researchers must establish common test and analytic norms to evaluate their algorithms, create interdisciplinary research teams of clinical cardiologists and machine learning experts together, obtain data sharing agreements between institutions that will allow for a large scale study needed for validation, and think about making their predictions actionable and meaningful through implementation tools that include software, education, and workflow, so that the research and implementation of machine learning can be realized in clinical practice to decrease cardiovascular disease and improve overall quality of care.

References:

[1] T. Liu, A. Krentz, L. Lu, and V. Curcin, “Machine learning based prediction models for cardiovascular disease risk using electronic health records data: systematic review and meta-analysis,” *European Heart Journal - Digital Health*, vol. 6, no. 1, pp. 7–22, Jan. 2025, doi: 10.1093/EHJDH/ZTAE080.

[2] S. Hossain, M. K. Hasan, M. O. Faruk, N. Aktar, R. Hossain, and K. Hossain, “Machine learning approach for predicting cardiovascular disease in Bangladesh: evidence from a cross-sectional study in 2023,” *BMC Cardiovasc Disord*, vol. 24, no. 1, pp. 1–28, Dec. 2024, doi: 10.1186/S12872-024-03883-2/TABLES/7.

[3] S. Y. Cho *et al.*, “Pre-existing and machine learning-based models for cardiovascular risk prediction,” *Sci Rep*, vol. 11, no. 1, p. 8886, Apr. 2021, doi: 10.1038/S41598-021-88257-W;SUBJMETA=1748,2099,4019,459,592,692,700,75;KWRD=DYSLIPIDAEMIAS,PREVENTIVE+MEDICINE.

[4] Y. Cai *et al.*, “Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: a systematic review,” *BMC Medicine 2024 22:1*, vol. 22, no. 1, pp. 1–18, Feb. 2024, doi: 10.1186/S12916-024-03273-7.

[5] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, “Can machine-learning improve cardiovascular risk prediction using routine clinical data?,” *PLoS One*, vol. 12, no. 4, p. e0174944, Apr. 2017, doi: 10.1371/JOURNAL.PONE.0174944.

[6] S. Subramani *et al.*, “Cardiovascular diseases prediction by machine learning incorporation with deep learning,” *Front Med (Lausanne)*, vol. 10, p. 1150933, Apr. 2023, doi: 10.3389/FMED.2023.1150933/BIBTEX.

[7] C. Krittanawong *et al.*, “Machine learning prediction in cardiovascular diseases: a meta-analysis,” *Sci Rep*, vol. 10, no. 1, pp. 1–11, Dec. 2020, doi: 10.1038/S41598-020-72685-1;SUBJMETA=114,1305,4019,592,631,692,75;KWRD=CARDIOVASCULAR+DISEASES,COMPUTATIONAL+BIOLOGY+AND+BIOINFORMATICS,MACHINE+LEARNING.

[8] D. R. Deepa, V. B. Sadu, C. G. Prashant, and D. A. Sivasamy, “Early prediction of cardiovascular disease using machine learning: Unveiling risk factors from health records,” *AIP Adv*, vol. 14, no. 3, Mar. 2024, doi: 10.1063/5.0191990/3279524.

[9] S. Dalal *et al.*, “Application of Machine Learning for Cardiovascular Disease Risk Prediction,” *Comput Intell Neurosci*, vol. 2023, no. 1, p. 9418666, Jan. 2023, doi: 10.1155/2023/9418666.

[10] K. Sumwiza, C. Twizere, G. Rushingabigwi, P. Bakunzibake, and P. Bamurigire, “Enhanced cardiovascular disease prediction model using random forest algorithm,” *Inform Med Unlocked*, vol. 41, p. 101316, Jan. 2023, doi: 10.1016/J.IMU.2023.101316.

[11] Y. Rimal, N. Sharma, S. Paudel, A. Alsadoon, M. P. Koirala, and S. Gill, “Comparative analysis of heart disease prediction using logistic regression, SVM, KNN, and random forest with cross-validation for improved accuracy,” *Scientific Reports 2025 15:1*, vol. 15, no. 1, pp. 1–14, Apr. 2025, doi: 10.1038/s41598-025-93675-1.

[12] M. D. Teja and G. M. Rayalu, “Optimizing heart disease diagnosis with advanced machine learning models: a comparison of predictive performance,” *BMC Cardiovasc Disord*, vol. 25, no. 1, pp. 1–12, Dec. 2025, doi: 10.1186/S12872-025-04627-6/TABLES/6.

[13] S. Mohan, C. Thirumalai, and G. Srivastava, “Effective heart disease prediction using hybrid machine learning techniques,” *IEEE Access*, vol. 7, pp. 81542–81554, 2019, doi: 10.1109/ACCESS.2019.2923707.