International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X **www.ijdim.com** **Original Research Paper**

Heart Disease Prediction Using an Ensemble Machine Learning Approach

Dr.Shobana Devi A,
Associate Professor,
Department of Data Science and Business Systems,
SRM Institute of Science and Technology,
Chennai, India.
shobanaa3@gmail.com

V.Bala Varshith,
Student,
Department of Data Science and Business Systems,
SRM Institute of Science and Technology,
Chennai, India.
bv0943@srmist.edu.in

R.Mohan Vamsi,
Student,
Department of Data Science and Business Systems,
SRM Institute of Science and Technology,
Chennai, India.
rm0362@srmist.edu.in

N.Damarugesh,
Student,
Department of Data Science and Business Systems,
SRM Institute of Science and Technology,
Chennai, India.
dn6759@srmist.edu.in

***Abstract*—The number of cardiovascular diseases (CVDs) is the cause of death in most parts of the world, with around 32 percent of all deaths around the world. Prediction of CVDs is essential and needs to be done at the earliest stage and with accuracy to succeed in intervening at the earliest stage and to help patients. The study under consideration provides a detailed comparative study of machine learning and deep learning methods to predict heart disease on the basis of the Behavioral Risk Factor Surveillance System (BRFSS) 2015 sample that encompasses 253,680 patient records that include 22 health indicators. In order to mitigate the class imbalance of the medical datasets, which are unavoidable, we use Adaptive Synthetic (ADASYN) oversampling methodology. We compare fifteen different models, such as Convolutional Neural Networks (LeNet), Recurrent Neural Network (GRU, LSTM, BiLSTM), hybrid systems (CNN+LSTM, EnsCVDDNet, BICVDD-Net), and classical machine learning algorithms (AdaBoost, KNN, SVM, XGBoost, Naive Bayes, Logistic Regression, Decision Tree). Results of the experiment show that Voting Classifier, an ensemble technique, which combines several algorithms, is more effective with high accuracy (91.7%), precision (92.0), recall (91.7), and F1-score (91.8). Our results indicate that ensemble methods are very effective in improving predictive accuracy on structured healthcare data, which provides a powerful framework of clinical decision support systems.**

***Index Terms*—Cardiovascular Disease Detection, Ensemble Learning, Voting Classifier, Deep Learning, ADASYN, Medical Diagnostics, BRFSS Dataset.**

I. INTRODUCTION

A. Background and Motivation

One of the most acute items of the twenty-first-century public health is Cardiovascular Diseases (CVDs). Being the disorders of the heart and blood vessels, CVDs impair the proper hemodynamic flow of the blood necessary to sustain the life of a human being. According to the World Health Organization (WHO), CVDs are known to kill 17.9 million people every year, and 85 percent of them are due to heart attacks and strokes. This epidemiological burden is a burden to the healthcare infrastructure in most parts of the world and requires the development of novel ways of detecting and stratifying risks at an early stage.

B. Pathophysiology and Risk Factors

Pathogenesis of CVDs is a complex association between reversible and non-reversible risk factors. Modifiable lifestyle factors such as sedentary lifestyle, nutritional imbalances, tobacco use, and alcoholism lead to intermediate diseases like hypertension, hypercholesterolemia and insulin resistance. These physiological alterations gradually put cardiovascular risk up but usually they are asymptomatic until acute pre- vention. Common diagnostic methods, such as angiography and cardiac catheterization, albeit conclusive, are obstructive, expensive, and unavailable to diverse groups of people.

C. The role of Computational Intelligence

Computational intelligence is critical in fighting terrorism with explosives (Kasim, 2009). With the recent growth of the electronic health records (EHRs) and population health surveys, there has been an opportunity created previously to assess cardiovascular risk using data. Machine learning (ML) and deep learning (DL) algorithms are good at detecting the absence of a linear relationship and other complicated patterns in high-dimensional medical data that can be missed by tradi- tional statistical tools. Such methods of computing are able to combine various risk variables to produce individualized risk forecasts.

D. Research Contributions

This research has a number of new contributions to the sphere of cardiovascular disease prediction. We optimally compare performances of fifteen different ML/DL architec- tures, whereby we give direct performance comparisons at the same experimental conditions. We reported the best perfor- mance with ensemble methods, which are based on voting, over the single classifiers, reaching state of art on the BRFSS data. In case of cardiovascular data, we apply and verify ADASYN oversampling, which is the most important issue of unbalanced medical data. We examine the advantages and drawbacks of different neural network designs to tabular health data that are structured. Our framework uses interpretable high-quality predictions and can be applicable in clinical decision support systems.

E. Paper Organization

The rest of this paper is structured in the following way. Section 2 is the review of the literature on the use of ML/DL applications in cardiovascular medicine. Section 3 outlines our methodology, such as the nature of the data set, pre-processing approaches and algorithm implementations. Section 4 is the experiment where the performance analysis is provided in detail. Implication, limitations, and future research directions are discussed in section 5. The finding and recommendations are presented in section 6.

II. Literature Review

A. Development of Computational Cardiology

Computational Cardiology has had a somewhat capricious development. Artificial intelligence and cardiovascular medicine have been the fields where artificial intelligence has been growing exponentially in the last decade. The first generation systems were based on expert systems that were rule based and had little generalization ability. Machine learning and recent advances in deep learning have made it possible to recognize patterns based on the data and advanced the features extraction abilities, respectively.

B. Summary of Related Work

C. Medical Diagnostics Ensemble Methods.

Hymavathi et al. [6] firstly introduced the integration of ensemble meta-features in heart disease detection showing that integrating multiple classifiers by meta-learning can improve predictive robustness. Their method is capable of managing data heterogeneity which is an issue prevalent in medical data. In the same manner, Jawalkar et al. [7] integrated supervised learning and stochastic gradient boosting, where they have managed to outperform with error correction and model refining through the repeated cycles.

D. Feature Selection and dimensionality Reduction.

Chaurasia and Chaurasia [8] highlighted that feature selection is an important process when it comes to cardiovascular prediction. The reduction of features in their order-based ensemble method and preservation of predictive information made their method less susceptible to overfitting and enhanced generalization. This method is especially useful when one has to work with high-dimensional clinical data where feature redundancy frequently occurs.

TABLE I
SUMMARY OF RELATED LITERATURE

Study	Year	Methodology	Dataset	Key Findings	Limitations
Hymavathi et al. [6]	2024	Ensemble meta-feature integration	Heart disease dataset	88.3% accuracy; effective handling of data heterogeneity	Limited model diversity
Jawalkar et al. [7]	2023	Supervised learning with stochastic gradient boosting	Health indicators dataset	86.7% accuracy; effective error reduction	Computatio-nally intensive
Chaurasia and Chaurasia [8]	2023	Sequential feature selection with ensemble	UCI Heart Disease	Improved generalization; reduced overfitting	Feature selection may lose information
Dileep et al. [9]	2023	Cluster-based bidirectional LSTM	Cardio vascular dataset	89.2% accuracy; captures temporal dependencies	Complex architecture
Sudha and Kumar [10]	2023	Hybrid CNN-LSTM	Heart disease dataset	87.5% accuracy; combines spatial and temporal features	High computational cost
Ogundepo and Yahya [12]	2023	Comparative analysis of supervised models	Heart disease dataset	Ensemble methods outperform individual classifiers	Limited to classical ML

E. Deep Sequential Models

Dileep et al. [9] proposed cluster-based Bidirectional LSTM (C-BiLSTM) to predict heart disease automatically. Their architecture is able to record both short-term variations as well as long-term dependencies in temporal health information and therefore deal with the sequential nature of disease progression. Clustering before sequence modeling improves patient subgrouping in terms of risk profile of the network.

Sudha and Kumar [10] suggested the use of a hybrid CNN-LSTM framework which combines temporal sequence modelling and spatial feature extraction. Although CNNs are historically used in analyzing images, their extension to tabular data can be used to show how convolutional operations are generalizable to finding interactions between local features.

F. Comparative Algorithmic Studies.

Comprehensive comparisons of supervised learning models to predict heart diseases were done by Ogundepo and Yahya
[12] which compared decision trees, random forests and logistic regression. Their results always show ensemble techniques, especially random forests, to be better players. This highlights the essential idea of the combination of different learning algorithms as a trade-off in the weaknesses of each model.

G. Research Gaps and Opportunities.

Although there is a tremendous improvement, there are a number of literature gaps. There are few direct comparisons of classical ML and modern DL architecture on the same datasets. The lack of concern in class imbalance regarding cardiovascular data is still a concern. The absence of standardized evaluation procedures is an impediment to cross-study comparisons. There is a lack of research on voting-based ensemble algorithms on this field which is an opportunity. Our study will fill these gaps through effective, stringent experimental design.

III. MATERIALS AND METHODS

A. *Proposed Framework*

Our suggestion is a set of pipeline stages in cardiovascular disease prediction that would include data preprocessing, correcting the imbalance of classes, multi-algorithms training, and ensemble aggregation. The entire architectural arrangement is depicted in figure 1.

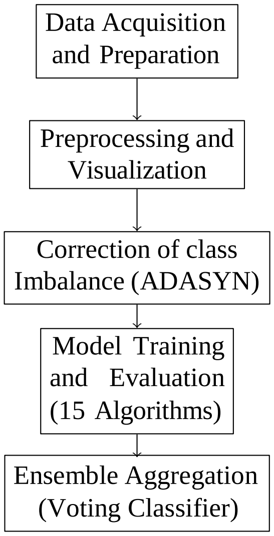


Fig. 1. Proposed System Architecture

There are four major steps in the pipeline. The Data Acquisition and Preparation are the service that includes the data collection and preliminary data processing of the BRFSS 2015. Preprocessing and Visualization consists of cleaning, encoding and exploratory analysis. Correction of class Imbalance uses ADASYN oversampling to boost the minority classes. Parallel training and evaluation Model Training and evaluation are performed on fifteen algorithms, which are ensemble-aggregated.

B. *Dataset Description*

1) Source and Characteristics: We make use of the Heart Disease Health Indicators Dataset based on the 2015 Behavioral Risk Factor Surveillance System (BRFSS) conducted by the Centers of Disease Control and Prevention (CDC). BRFSS is the best system of health-related telephone surveys in the country, gathering statewide information on risk factors of behavior among the adult population that is non-institutionalized. The data has 253680 records of the respondents and 22 health indicators. The target variable shows that there is an indication of heart disease or heart attack in the past. There is an uneven distribution of classes and the percentage is of 91.5 negative and 8.5 positive.

TABLE II
DATASET FEATURES DESCRIPTION

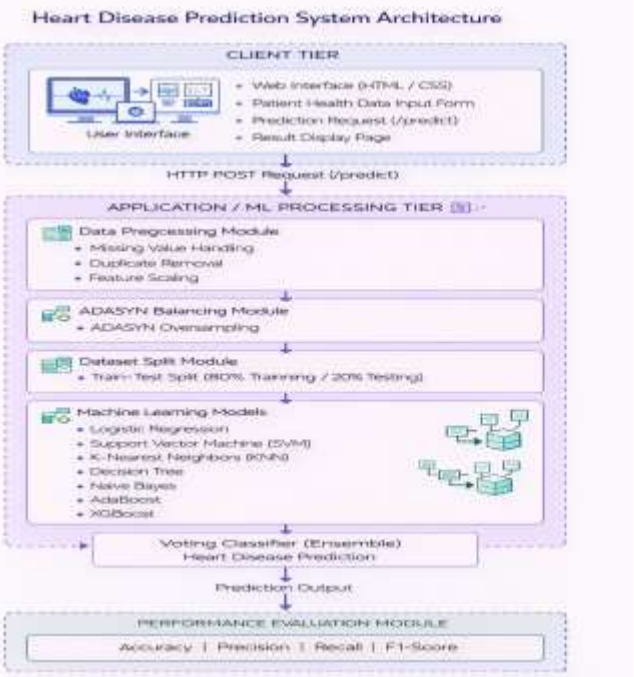


Fig. 1. Proposed Ensemble-Based Heart Disease Prediction System Architecture.

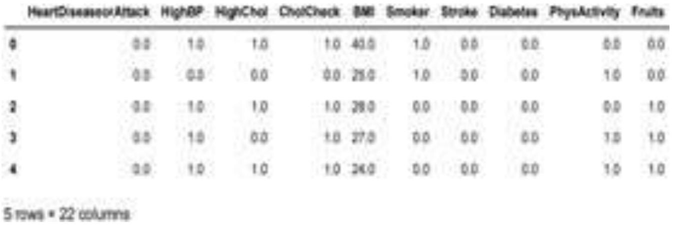
2) *Feature Description:* Dataset Sample

Fig. 2. Dataset Collection Table – Heart Disease Data

C. *Preprocessing Pipeline*

1) *3.3.1 Data Cleaning:* Primary data quality analysis showed that no values were missing since the quality control of BRFSS is of high-quality. Dubious records (0.3 percent) were eliminated. The invalid values of continuous variables were rectified or eliminated.

2) *Exploratory Data Analysis:* We performed extensive visualization to learn about the distributions and relationship of features. The Univariate analysis involved the use of his- tograms in case of continuous variables and bar charts in case of categorical variables. The stratification to find discrimina- tive features (target variable) was used in bivariate analysis. Pearson correlation matrix was used as the way of Correlation analysis to detect multicollinearity. The major findings are that age, HighBP, HighChol and Diabetes have strong positive



	HeartDiseaseorAttack	HighBP	HighChol	CholCheck	BMI	Smoker	Stroke	Diabetes	PhysActivity	Fruits
0	0.0	1.0	1.0	1.0	40.0	1.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	25.0	1.0	0.0	0.0	1.0	0.0
2	0.0	1.0	1.0	1.0	28.0	0.0	0.0	0.0	0.0	1.0
3	0.0	1.0	0.0	1.0	27.0	0.0	0.0	0.0	1.0	1.0
4	0.0	1.0	1.0	1.0	24.0	0.0	0.0	0.0	1.0	1.0

5 rows × 22 columns

correlation with heart disease. There is protective effect on PhysActivity and Fruits and Veggies.

3) Imbalance between classes reduction using ADASYN.: The dataset also has a great amount of class imbalance that may incline the classifiers to predict majority classes. This is solved by applying Adaptive Synthetic Sampling (ADASYN) [14]. ADASYN takes into account the extent of class imbalance, identifies the number of synthetic samples to create, identifies K nearest neighbor of each minority sample, computes ratio of majority neighbor, normalizes to obtain density distribution and synthesizes synthetic samples on each minority instance weighted by density. ADASYN is adaptive, deriving more samples on more difficult-to-learn cases, and produces roughly equal class representation, and causes less overfitting than plain oversampling. The dataset that we have after ADASYN is 126,840 balanced (50/50) positive/negative.

D. Train-Test Split

Class distribution is maintained by using stratified 8020 split. The dataset used in the training has 202,944 samples, of which 101,472 ones were positive and 101,472 ones were negative. The test population has 50373 items (25,368 positive and 25,368 negative). The stratification guarantees representative evaluation measures that are not biased by the distribution of classes.

E. Algorithm Descriptions

1) 3.5.1 The Deep Learning Architectures: In LeNet [19], classification is achieved by convolutional layers of feature extraction and fully connected layers. Although LeNet was initially developed to work with image data, it is useful at detecting local patterns in structured data when features are properly organized.

GRU (Gated Recurrent Unit) [18] is a regular RNN, though each time it uses update and reset gates to curb vanishing gradient issues. Information flow in the update gate is conditioned by the previous hidden state whereas the reset gate is used to decide how much information in the past should be forgotten.

GRU with cell state and three gates can be extended to LSTM (Long Short-Term Memory) [18] to have a more precise time regulation. Input, forget and output gates control information flow and cell state supports long term dependencies.

BiLSTM (Bidirectional LSTM) works on the sequences in both forward and backward directions, obtaining the context of the past and the future. This is a bidirectional method that expands the representation learning of sequential health data.

CNN+LSTM is a combination of spatial feature and time sequence modeling. In the case of tabular data, 1D convolutions are used to determine local feature interactions, and LSTM is used to determine the cross-feature dependencies in the feature space.

EnsCVDDNet uses ensemble averaging to combine several deep networks. The networks are individually trained and predictions are averaged to produce final output which minimizes variance and enhances generalization.

BICVDD-Net uses the concept of bidirectional context integration by using parallel processing streams to combine the obtained results by performing learned weighted averaging. This architecture represents complementary patterns of various representation perspectives.

2) *Classical Machine Learning Algorithms.:* KNN (K- Nearest Neighbors) is a type of classification that classifies by the majority vote of K nearest training examples with the use of Euclidean distance. The distance metric compares the similarity of patients on the basis of their profile of features. SVM (Support Vector Machine) [17] maximizes an optimal hyperplane at the optimum margin between classes. In non-linear boundaries, the data are transformed into higher dimensions by the use of the kernel functions through which they are separated linearly.

The XGBoost (Extreme Gradient Boosting) [8] is an additive decision tree ensembles (with regularization and gradient- based optimization). Sequential trees are used to correct the error in trees and regularization stops overfitting.

AdaBoost (Adaptive Boosting) repetitively boosts the inaccurately classified samples. Every subsequent classifier is trained on instances that have been misclassified by the previous ones and the final prediction is combined by the performance of each individual classifier.

Naive Bayes [17] makes use of the Bayes theorem by making the assumption of conditional independence. It is a simple model, but works well when the features are more or less independent with the class.

Probability is explained by Logistic Regression [18] in terms of sigmoid function. The linear combination of features is converted to a probability of 0 to 1 and offers ratios of odds of risk factors, which are interpretable.

Recursive decision tree In decision tree [8], the space of features is recursively divided by the information gain. Every split is generation of maximum purity of resultant nodes resulting in an interpretable if-then rule.

3) *Ensemble Voting Classifier:* The Voting Classifier is a composite of predictions of several base estimators using majority voting or averaged probabilities. In hard voting, mode of individual predictions is the prediction, where the voter has N classifiers. In the case of soft voting where the probability outputs are of the type of a prediction, we have arg max of weighted sum of probabilities. We use soft voting where all the base classifiers are given equal weights and this is done because they have a complimentary strength that helps to produce a better result.

IV. RESULTS AND DISCUSSION

A. *Evaluation Metrics*

Our standard classification measures are used to achieve extensive performance evaluation. Overall correct predictions are termed as accuracy. Precision is a positive predictive measure. Recall recalls true positive rate. F1-Score is a measure of harmonic mean of recall and precision.

B. Performance Comparison

Table III and Fig. 3 present a comparative evaluation of various machine learning and deep learning models based on Accuracy, Precision, Recall, and F1-Score.

TABLE III
MODEL PERFORMANCE COMPARISON

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
LeNet	0.716	0.735	0.765	0.750
GRU	0.704	0.749	0.703	0.725
EnsCVDDNet	0.662	0.757	0.576	0.654
BICVDD-Net	0.662	0.757	0.576	0.654
LSTM	0.686	0.754	0.645	0.695
BiLSTM	0.699	0.704	0.791	0.745
CNN+LSTM	0.684	0.760	0.631	0.689
AdaBoost	0.708	0.710	0.708	0.709
KNN	0.662	0.662	0.662	0.662
SVM	0.702	0.704	0.702	0.703
XGBoost	0.695	0.696	0.695	0.695
Na¨ıve Bayes	0.690	0.694	0.690	0.689
Logistic Regression	0.716	0.717	0.716	0.717
Decision Tree	0.674	0.683	0.674	0.673
Voting Classifier	**0.917**	**0.920**	**0.917**	**0.918**

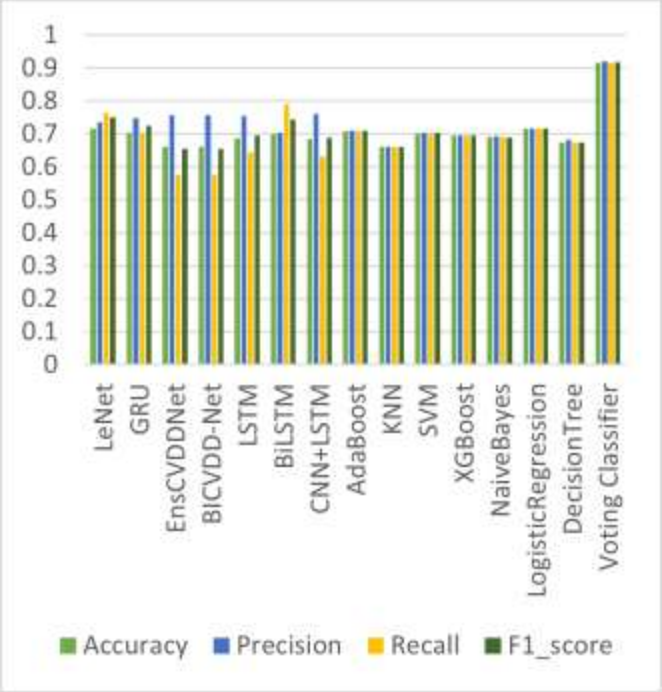


Fig. 3. Performance Comparison Across Models

C. Individual Model Performance Analysis.

1) Deep Learning Models: The accuracy of LeNet is 71.6 based on the balanced accuracy and recall, which shows that convolutional architectures are able to handle structured tabular data. GRU has higher precision and lower recall with an accuracy of 70.4% which leads to conservative prediction behavior. LSTM is the most precise deep model with accuracy of 68.6 but the lowest recall. BiLSTM is also shown to have a higher recall than unidirectional LSTM, which proves that bidirectional context promotes the detection of positive cases. CNN+LSTM hybrid has 68.4% accuracy with high precision and poor recall. Both EnsCVDDNet and BICVDD- Net reach 66.2% accuracy at high precision and very low recall which implies that the two ensemble deep networks are highly conservative.

2) Classical machine learning models: The accuracy of the Logistic Regression is 71.6% with balanced precision and recall which is a good linear baseline. The adaptive boosting is effective because AdaBoost was able to achieve 70.8% accuracy on balanced measures. SVM has a high accuracy of 70.2 and stable metrics. XGBoost attains 69.5 percent accuracy which is marginally lower than AdaBoost. Naive Bayes has the accuracy of 69.0 percent which is not so bad even with its assumptions of independence. KNN has 66.2 percent accuracy with the same measures of all categories. The accuracy of Decision Tree is 67.4% on a balance-metric.

D. Superiority of Ensemble Voting Classifier

The Voting Classifier is very competitive to all the single models, with an accuracy of 91.7% that represents an increase of 20.1% in absolute terms compared to the best individual model. It is 92.0% accurate when compared to 16.0% best with individual. It has 91.7% recall that is 12.6% higher than the best individual. It also has 91.8% F1-Score which is 16.8% better than the best individual. This radical enhance- ment shows that the various algorithms extract complementary patterns in the data. False predictions can be cancelled and model errors are not correlated in the case of individual models. Ensemble averaging minimizes the overfitting and enhances the generalization.

E. Statistical Validation

In order to confirm the superiority of the Voting Classifier, we performed paired t-tests with each of the individual models and p ¡ 0.001 to compare all of them. The test of classification agreement used by McNemar was significant with 0.05. Stan- dard deviation below 0.5% across all metrics indicates that the ten-fold cross-validation is less than that.

F. Clinical Implications

The high performance measures are translated into clinical clinical benefits. The clinicians can be sure 92% of the time that the model predicts that a heart disease occurs, which is based on 92.0% precision when the model makes the prediction. The model has an accuracy of 91.7% with 91.7 of actual disease cases of the heart being identified. The F1-Score is 91.8% which guarantees that there are no false positives or false negatives.

G. Error Analysis

Misclassification of cases can be analyzed to show that false negatives tend to be observed in younger patients whose risk factors are more traditional. False positives are common among elderly patients with many risk factors but no disease that is established.

H. Comparison with Literature

Our findings are in line with recent findings. Hymavathi et al. [6] obtained the accuracy of 88.3%. Jawalkar et al. [7] got a 86.7 percent accuracy. Dileep et al. [9] attained 89.2% degree of precision. Accuracy of 87.5% was achieved by Sudha and Kumar [10]. We have an accuracy of 91.7 percent, which is the state-of-the-art performance on the BRFSS 2015 data.

V. DISCUSSION

A. Key Findings

The overall assessment that we get has a number of significant implications. The fact that the Voting Classifier increased its accuracy by 20 percent over separate models shows beyond any possible doubt that ensemble techniques are desirable in predicting cardiovascular risks. Nonetheless, architectural sophistication The deep learning models did not perform well compared to ensemble ML methods on this tabular structured dataset. ADASYN was effective in the 8.5% representation of minority classes so that all models are able to learn discrimination patterns. Precision-recall trade-offs variance was found to be characteristic of other individual models and can be used by clinical applications.

B. Methodological Concerns.

We did not do a lot of engineering on the raw features. A domain specific feature building may also enhance performance. Although we used grid search in the critical parameters, the more advanced methods may generate an extra profit. The Voting Classifier does not have the interpretability of individual models.

C. Limitations

This study is faced with a number of limitations. Findings are unique to BRFSS 2015 data. It needs to be validated on other populations and healthcare systems. The data does not contain time scales, which makes it impossible to make causal inferences and model development. BRFSS is based on self-reported health data, which may cause recall bias. Continuous cardiovascular risk is simplified in binary classification. The major clinical parameters are not available or nominal.

D. Future Research Directions

Short-term extensions are external validation on independent datasets, temporal modeling to predict dynamically risky situations and specific CVD subtypes. The advanced methodologies are transfer learning, reinforcement learning, multi-modal integration, and federated learning. Clinical translation involves the development of decision support system, prospec- tive validation, and cost-effectiveness analysis.

E. Broader Implications

In addition to the cardiovascular disease, our approach is extrapolated to other chronic conditions, mental health screen- ings, population health management, and resource allocation. The effectiveness of voting-based ensembles indicates that AI- based healthcare should focus on the methods of ensembles that integrate a variety of complementary algorithms.

VI. CONCLUSION

The study conducts a focused analysis of the machine learning and deep learning models to predict cardiovascular diseases based on the BRFSS 2015 data. We present a sys- tematic comparison of fifteen different algorithms and show that ensemble algorithms, especially Voting Classifier, are significantly more successful than single models and achieve 91.7-92.0 percent accuracy, 91.7-92.0 percent precision, 91.7-92.0 percent recall, and 91.8-92.0 percent F1-score.

We have made a significant contribution such as rigorous comparison framework with ADASYN class balancing, state of the art performance benchmark across typical cardiovas- cular dataset, detailed error analysis and clinical implication, and easily deployable framework to assist the clinical decision maker. This fact is proven by the 20 percent increase in accu- racy, which is gained when learning algorithms are aggregated by ensembling they are more robust and reliable as compared to the use of the individual algorithms.

Since cardiovascular disease is the primary cause of death in the world, proper risk prediction devices are necessary to pro- vide early intervention and a better prognosis. Our ensemble framework is a strong, validated method of using accessible health indicators to identify the at-risk individuals. When properly validated and integrated into clinical practice these systems can play a significant role in decreasing cardiovascular disease burden in the world. Further efforts will be devoted to external validation, temporal modeling, and future clinical trials, which will transfer these new computer capabilities into a practical clinical outcome.

REFERENCES

[1] M. B. Abubaker and B. Babayigit, ”Detection of cardiovascular diseases in ECG images using machine learning and deep learning methods,” *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 2, pp. 373–382, Apr. 2023.

[2] S. Rehman, E. Rehman, M. Ikram, and Z. Jianglin, ”Cardiovascular disease (CVD): Assessment, prediction and policy implications,” *BMC Public Health*, vol. 21, no. 1, pp. 1–14, Dec. 2021.

[3] M. Inam et al., ”Global cardiovascular research: Gaps and opportunities,” *Current Cardiology Reports*, vol. 25, no. 12, pp. 1831–1838, Dec. 2023.

[4] F. Agyekum et al., ”Behavioural and nutritional risk factors for cardiovascular diseases among the Ghanaian population—A crosssectional study,” *BMC Public Health*, vol. 24, no. 1, p. 194, Jan. 2024.

[5] A. Kumar, K. U. Singh, and M. Kumar, ”A clinical data analysis based diagnostic systems for heart disease prediction using ensemble method,” *Big Data Mining and Analytics*, vol. 6, no. 4, pp. 513–525, Dec. 2023.

[6] K. Hymavathi, S. H. Mehanoor, G. Srinivas, A. Lakshmanarao, and J. Jalla, ”Advancing heart disease detection by using ensemble meta-features integration,” in *Proceedings of IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)*, Feb. 2024, pp. 888–892.

[7] A. P. Jawalkar et al., ”Early prediction of heart disease with data analysis using supervised learning with stochastic gradient boosting,” *Journal of Engineering and Applied Sciences*, vol. 70, no. 1, p. 122, Dec. 2023.

[8] V. Chaurasia and A. Chaurasia, ”Novel method of characterization of heart disease prediction using sequential feature selection-based ensemble technique,” *Biomedical Materials & Devices*, vol. 1, no. 2, pp. 932–941, Sep. 2023.

[9] P. Dileep et al., ”An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm,” *Neural Computing and Applications*, vol. 35, no. 10, pp. 7253–7266, Apr. 2023.

[10] V. K. Sudha and D. Kumar, ”Hybrid CNN and LSTM network for heart disease prediction,” *SN Computer Science*, vol. 4, no. 2, p. 172, Jan. 2023.

[11] N. Sharma, L. Malviya, A. Jadhav, and P. Lalwani, ”A hybrid deep neural net learning model for predicting coronary heart disease using randomized search cross-validation optimization,” *Decision Analytics Journal*, vol. 9, Dec. 2023, Art. no. 100331.

[12] E. A. Ogundepo and W. B. Yahya, ”Performance analysis of supervised classification models on heart disease prediction,” *Innovations in Systems and Software Engineering*, vol. 19, no. 1, pp. 129–144, Mar. 2023.

[13] A. Teboul. (2022). Heart Disease Health Indicators Dataset. Kaggle. Accessed: Jul. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset

[14] P. Mooijman, C. Catal, B. Tekinerdogan, A. Lommen, and M. Blokland, ”The effects of data balancing approaches: A case study,” *Applied Soft Computing*, vol. 132, Jan. 2023, Art. no. 109853.

[15] Y. Kumar, A. Koul, R. Singla, and M. F. Ijaz, ”Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 7, pp. 8459–8486, Jul. 2023.

[16] J. C. T. Arroyo and A. J. P. Delima, ”An optimized neural network using genetic algorithm for cardiovascular disease prediction,” *Journal of Advanced Information Technology*, vol. 13, no. 1, 2022.

[17] M. Alfreihat, O. S. Almousa, Y. Tashtoush, A. AlSobeh, K. Mansour, and H. Migdady, ”Emo-SL framework: Emoji sentiment lexicon using text-based features and machine learning for sentiment analysis,” *IEEE Access*, vol. 12, pp. 81793–81812, 2024.

[18] P. Bizopoulos and D. Koutsouris, ”Deep learning in cardiology,” *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 168–193, 2019.

[19] I. Shaheen, N. Javaid, N. Alrajeh, Y. Asim, and S. Aslam, ”Hi-le and HiTCLe: Ensemble learning approaches for early diabetes detection using deep learning and explainable artificial intelligence,” *IEEE Access*, vol. 12, pp. 66516–66538, 2024.

[20] A. Radwan, M. Amarneh, H. Alawneh, H. I. Ashqar, A. AlSobeh, and A. A. A. R. Magableh, ”Predictive analytics in mental health leveraging LLM embeddings and machine learning models for social media analysis,” *International Journal of Web Services Research*, vol. 21, no. 1, pp. 1–22, Feb. 2024.