International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Cutting-Edge Neural Network for Early Cardiovascular Disease Prevention

Shivganga Udhan[1], Bankat Patil[2]

Submitted: 21/08/2023 **Revised**: 06/10/2023 **Accepted**: 21/10/2023

Abstract: Cardiovascular diseases (CVD) are common and life-threatening, requiring early detection to reduce mortality. This study presents an efficient system for predicting and preventing CVD. It utilizes a hybrid dataset from various sources datasets, preprocesses them. Feature selection methods like ANOVA and CHI2 enhance prediction accuracy. The Class Balanced Feature Selection Deep Neural Network model is an enhanced deep neural network that incorporates balanced data and utilizes a prominent feature selection technique. Multiple classifiers, including Naïve Bayes, Decision Tree, CBFS-DNN, SVM, Random Forest, and KNN are trained on the hybrid dataset using the selection of features and class balancing. The DNN with CHI2 selection achieves an impressive 99.79% accuracy, demonstrating high Precision, Recall, and 97.3%, 97.2%, and 97.2%, respectively, for the F1 Score.

Keywords: *Cardiovascular diseases, Machine Learning, Deep Neural Network, Prediction, and Prevention.*

1. Introduction

CVD has surpassed cancer as the top cause of mortality worldwide by accounting for one out of every three fatalities. In particular, in developing nations such as India, CVD has become a major cause of death, with the World Heart Federation estimating that CVD accounts for 26% of all deaths. [1]. The World Health Organization (WHO) predicted that by 2030, this number might reach 35% of all world deaths. Cardiovascular disease (CVD) disproportionately affects developing and underdeveloped countries, primarily because of restricted access to healthcare facilities and emergency medical services, as well as harmful lifestyle practices such as tobacco use, excessive alcohol consumption, and smoking.

However, concerning aspect is not just the overall number of CVD-related deaths but also the age at which people are experiencing these conditions. There is a growing alarm as CVD is no longer restricted to older age groups but is increasingly observed across a broader age range, spanning from 25 years to older individuals. This trend is attributed to modern lifestyles and dietary habits.

In India, CVD affects 52% of the population before the age of 70, while in Western countries, this percentage is notably lower at 23%. Addressing cardiovascular risk factors and implementing preventive measures are crucial in reducing the global burden of this disease.[2].

Every year, cardiovascular disease (CVD) claims more lives than any other cause, with a staggering death toll of approximately 17.9 million, accounting for 31% of global mortality. Of these deaths, a substantial 85% are attributed to stroke and heart attacks. According to the WHO, it is projected that nearly 23.6 million deaths, primarily from stroke and heart disease, will be attributable to CVDs by the year 2030.[3]. The data depicted in Figure 1 underscores that CVD continues to be the leading cause of death to this day.

[1]*Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University. Aurangabad, India.*
[2]*Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University. Aurangabad, India.*
Corresponding Author: Shivganga Udhan

Email: shivganga168@gmail.com

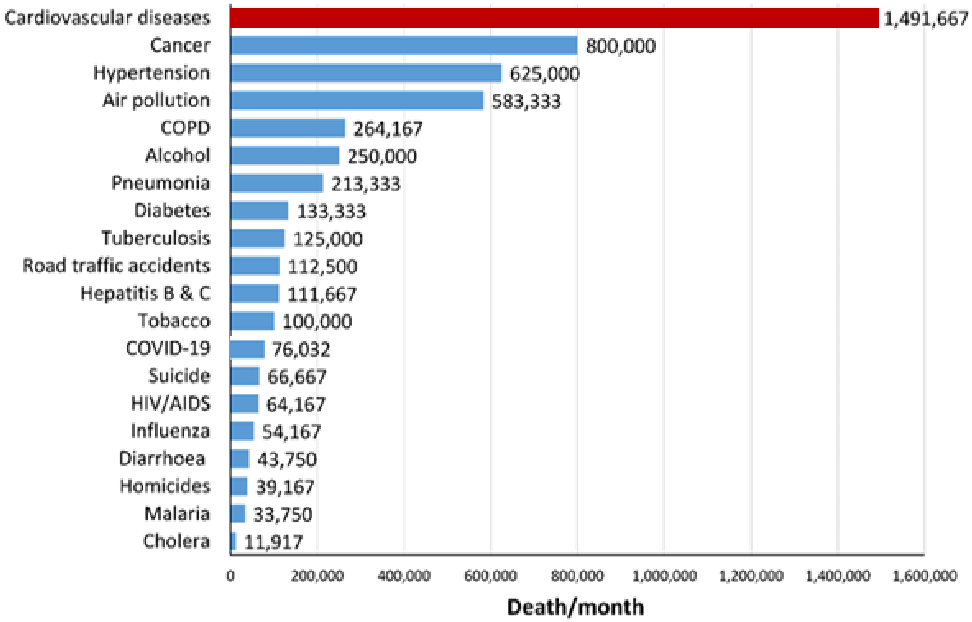


Fig 1: The mortality rate of the world's 20 major causes of death [3].

Preventing CVD disorders is achievable through the management of key factors, such as unhealthy dietary choices, unbalanced lifestyles, physical inactivity, and excessive alcohol consumption. It is also critical to detect those individuals having significant risk of CVD illnesses. Therefore, accurate diagnosis plays a crucial role, allowing for timely intervention and effective treatment to mitigate the risk of disease progression and severity.[4]

Diagnosing and treating patients in the early stages of cardiovascular disease remains a significant challenge for cardiologists. The conventional method involves identifying diseases through the analysis of patient symptoms and medical history, including assessments such as, BP, ECG tests, blood sugar levels and cholesterol profiles.[5]

The traditional approach is both time-consuming and costly, but it can be streamlined and made more efficient through the application of ML and DL algorithms. This streamlined process not only saves considerable time but also enhances diagnostic efficiency. Consequently, our research has produced DL models designed to classify heart disease effectively. Particularly, the use of both ML and DL has significantly enhanced the ability to diagnose clinical data.[6]

Technologies based on big data are being implemented by hospitals gradually as the amount of data in the world keeps developing. Utilizing data analysis in the medical field is yielding substantial advantages. The effective integration of ML and DL techniques is not only improving the efficiency of healthcare providers but also boosting the overall productivity of healthcare services. [7]

Machine learning empowers computers to perform tasks more efficiently, akin to human capabilities.[8]. Machine learning gives systems the ability to autonomously access and accumulate knowledge, improving their capabilities over time. This technology has streamlined various aspects of life and has become an indispensable tool across multiple industries, including agriculture. [9], banking [10], optimization [11], robotics [12], structural health monitoring [13], and many more. Its applications span from object detection, image analysis, color and pattern recognition in cameras to data organization, sorting, and even audio-to-text transcription.

Early cardiovascular disease forecasting methods played a pivotal role in guiding decisions to implement necessary changes in high-risk patients, ultimately leading to a significant reduction in their risk factors. Researchers worldwide are placing significant emphasis on employing machine learning algorithms within the healthcare sector to predict various diseases. The efficiency in medical procedures has significantly increased as a result of the use of ML in this field.[7][14]

Remaining sections of the paper follow this organization: Section 2 encompasses an in-depth literature survey, Section 3 details the methodologies employed, Section 4 centers on presenting and discussing the results, and the research findings have been determined in Section 5.

2. Literature Review

With their massive approaches, ML and DL have a study scope that would aid in the prediction of the CVD priory and detect their behavioural patterns in vast volumes of data. The consequences of these predictions will help healthcare practitioners make educated judgements and make early diagnoses, lowering the likelihood of patients suffering life-threatening conditions. Various ML and DL models are used for prediction of CVD are studied and compared in our research.

Using the UCI data set, Lama A. Alqahtani et al. explored a range of machine learning techniques, including NB, MLP, RF) and DT to forecast the likelihood of cardiac disease. During the data preparation stage, the feature selection was done using a brute-force method, and the parameter optimization was done using a grid search mechanism. According to the trials, Random Forest also showed the best with a precision of 0.93 and AUC of 0.95, respectively[3].

This study introduces the FWAFE-ANN and FWAFE-DNN hybrid diagnostic algorithms and assesses their performance using data from the Cleveland online cardiac disease database. According to the Matthews correlation coefficient (MCC), receiver operating characteristics curve, accuracy, sensitivity, and specificity, the proposed approaches are evaluated. The experimental findings show that our models outperform eighteen previously suggested techniques, which had accuracy ranges of 50.00% to 91.83%. The efficiency of our models is also superior to current state-of-the-art ML techniques for diagnosing heart illness.[15].

This study investigates machine learning methods for predicting cardiac failure. It assesses various algorithms, including KNN, SVM, DT, RF, and LR, aiming to enhance accuracy. Boosting methods like XGBoost and CatBoost are also applied, along with data visualization. The top performers are DT and RF with 95% accuracy, followed by KNN at 93.33%, while XGBoost achieves 91.67%. ANN, CatBoost and SVM reach 90%, and LR achieves 88.33% accuracy.[16].

Utilizing machine learning and health information from patients, this study improves the early-stage heart failure identification process. To allow users choose important characteristics and boost performance, it introduces a Narrative Principal Component Heart Failure approach for feature engineering. A new feature set was developed through optimization, reaching the best accuracy. Extensive tests were performed on the dataset, which was based on eight best-fit characteristics. The proposed decision tree method outshone other models and achieved an impressive 100% accuracy, surpassing previous studies. This promises significant advancements in heart failure detection. [17].

The creation of a hybrid approach combining a genetic algorithm with a radial basis function is the main goal of this research, referred to as GA-RBF, to enhance the accuracy of coronary disease detection through feature selection. The system demonstrated an initial accuracy of 85.40% using 14 attributes. However, through attribute reduction, with just nine parameters, the predicted accuracy dramatically increased to 94.20 percent. This highlights the enhanced performance achieved by our approach.[18].

Table 1 below shows the Comparison of Existing System a survey for finding methodologies used by researchers in the similar domain. It also puts light on advantages results and research gaps for further enhancement. This will help to isentify the research gap in the existing system. The gap identified will be considered to decide the objectives and scope of the system.

Table 1: Comparison of Existing System

Reference	**Year**	**Methodologies**	**Advantages/ Results**	**Research gap**
Lama A. Alqahtani [19]	2020	Brute Force Technologies for feature selection & Grid Search Techniques for parameter optimization. NB, MLP,RF,DT techniques were applied for prediction.	RF achieved 93.00% accuracy	Need to be tested with more classifiers and only one dataset was used for research.
P. Swarnalatha [20]	2020	cluster-based DT learning (CDTL) system. RF was applied	RF 89.30%	Accuracy can be improved. Only 1 dataset was tested.
Eqbal Ahmad [1]	2020	Agile Methodology for project with machine learning model like SVM,RF,KNN,DT	KNN achieved 85.83% accuracy	Neural Network implementation is missing. Accuracy is on lower side

Ashir Javeed [15]	2020	Features are chosen using FWAFE, then ANN, then DNN, and finally the process is completed.	DNN with 91.83% accuracy	Single Dataset implementation and only 1 feature selection method
Pooja Rani [21]	2020	The data were selected by the study using a clever way, and they were then processed for analysis. Afterward, it employed a number of computer programs to generate predictions about anything.	RF achieves 86.6% accuracy	Only one Feature selection method was used. No Deep learning model was implemented. Accuracy is on the lower side.
Sibo Patro [22]	2021	SSA-NN, BO-SVM, NB and KNN	BO-SVM with 93.3% accuracy	Accuracy can be improved. Feature Selection Not Done.
M.S. Nawaz [23]	2021	Gradient Descent Optimization (GDO) and then SVM, KNN, RF, ANN were applied	GDO with 98.54% accuracy	Limited dataset. Neural network implementation is missing
D.P. Yadav [24]	2022	SVM, NB, RF, KNN with threefold cross validation	NB with a precision of 87.78%	One feature optimization method was employed. It is possible to boost accuracy.
Hafsa Binte Kibria [25]	2022	Random Over Sampler for balancing the dataset. weighted score fusion approach with ML model like ANN, SVM, LR, DT, RF, AdaBoost	AdaBoost with Multiclass achieved accuracy of 75% and Binary classification accuracy was 95%	Accuracy can be improved. Neural Network is missing.
P. Rahman [16]	2023	KNN, SVM, RF, DT, LR, XGB, ANN, CATBoost	Maximum accuracy, 95%, was attained by the DT and RF algorithms.	Other Optimization technique is not tried.
V. V. R. Karna [26]	2022	SVM,J48,KNN,AdaBoost M1, Bagging, RotF	Highest accuracy of 96.34% using KNN	No class balancing and other preprocessing is applied.

3. Methodologies

The resources and methods used in the investigation that is suggested are explained in detail in this section.

3.1. Data repository Description

Various datasets are studied and combined to get hybrid dataset. This integration allows for a wider range of factors to contribute to the CVD risk assessment, including genetic information, environmental factors, and lifestyle choices. In our research we have studied and analysed different CVD datasets cleveland dataset, Hungary dataset, switzerland dataset, statlog dataset and VA Long Beach Datasets.

Table 2: No of records from different datasets

Datasets	**No. of Records**
Cleveland	298
Hungary	294
Switzerland	123

VA Long Beach	205
Statlog	271
Total	1191

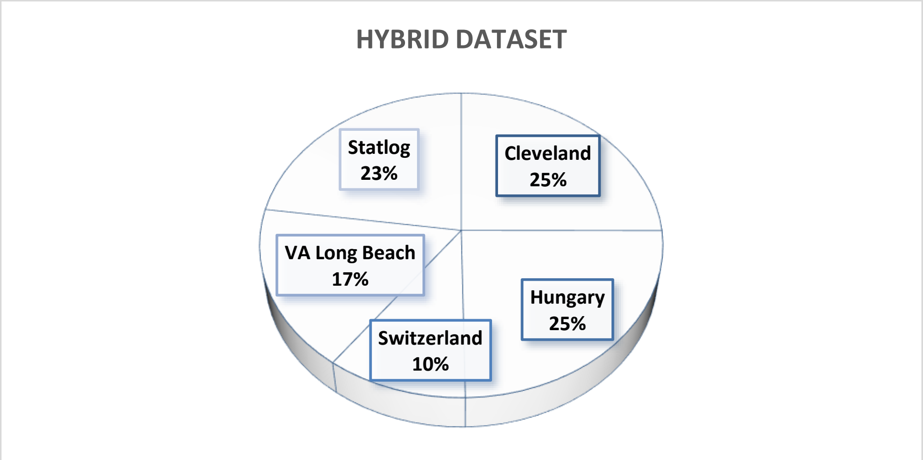


Fig 2: Hybrid dataset

3.2. System architecture

Figure 3 illustrates the sequential methodology of the proposed model. Initially, an analysis of various cardiovascular disease (CVD) datasets is conducted. Subsequently, datasets with shared features are merged to create a hybrid dataset. Following the creation of this hybrid dataset, preprocessing techniques are applied. In real-world scenarios, the number of healthy individuals often outweighs the number of patients with CVD. Imbalanced data can lead to biased models that may overlook important patterns in the minority class. To avoid this problem first of all the class balancing is done.

Missing data is handled so that final output is not biased because of missing data. To bring all features on a comparable scale and to make sure that each feature contributes equally to the learning process, Minmax and Standard scaling techniques are applied. Outlier handling is crucial to ensure accurate and reliable results in data analysis and modelling. It involves identifying and treating extreme values that can adversely impact statistical measures and machine learning algorithms. By selecting the most crucial element and eliminating redundant and irrelevant features, attribute selection improves ML and boosts the prediction power of ML algorithms. Different feature selection techniques are utilized by researchers such as Chi2 and ANOVA. This approach facilitates the identification of crucial features that enhance accuracy.

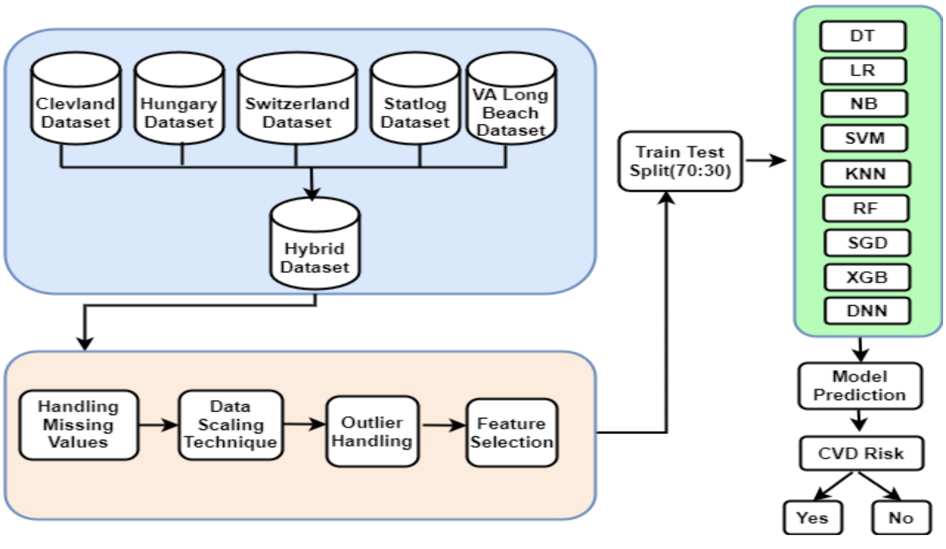


Fig 3: Architecture diagram for CVD risk prediction.

Following feature selection, 70 percent of the data exist for training, and 30% for testing. Various learning techniques, including SVM, NB, SGD, KNN, Logistic Regression, DT, RF, and XGBoost are applied to train and test the model for CVD risk prediction.

3.3. Classifier

This research harnesses technology derived from Artificial Neural Networks (ANN), specifically referred to as Deep Learning (DL). In a standard neural network, the core components consist of numerous interconnected processing units known as neurons. These neurons play a pivotal role in generating a series of real-valued activations, which collectively contribute to achieving the desired outcome. The mathematical model of an artificial neuron, also referred to as a processing component. Figure 4 shows a simplified architectural representation.

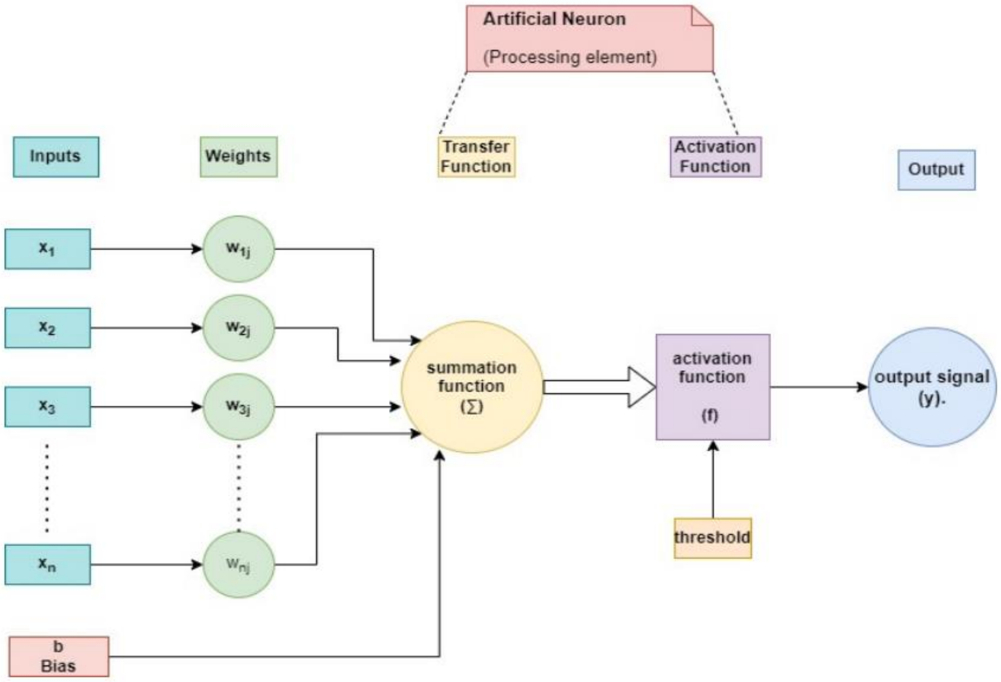


Fig 4: Mathematical model for Artificial Neuron [6]

All above models are used to check the performance of the system for detection of CVD. It has been observed that DNN performs better with respect to evaluation parameters such as precision, accuracy, F1 score, recall and specificity.

ANN which are intended to mimic the function of neurons in the human brain, are the foundation of deep learning. A perceptron, often referred to as an artificial neuron, simulates how an actual neuron operates by processing inputs with varying weights. A function is determined using these weighted inputs to produce an output. For instance, N inputs, one for each characteristic, are sent to the neuron. These inputs are combined by the neuron, which then activates them to produce the output depicted in Figure 5. An input's weight determines how important it is; the neural network gives inputs with higher weights a higher significance rating. [6].

To fine-tune the output of each individual perceptron, the neural network can adjust its bias parameter, ensuring the best possible fit between the model and the data. An activation function serves as the transformation between inputs and outputs, typically resulting in a threshold-based output. Various activation functions exist, including linear, logistic, identity, SoftMax, binary step, sigmoid, tanh, ReLU, and unit, each with distinct properties and applications.

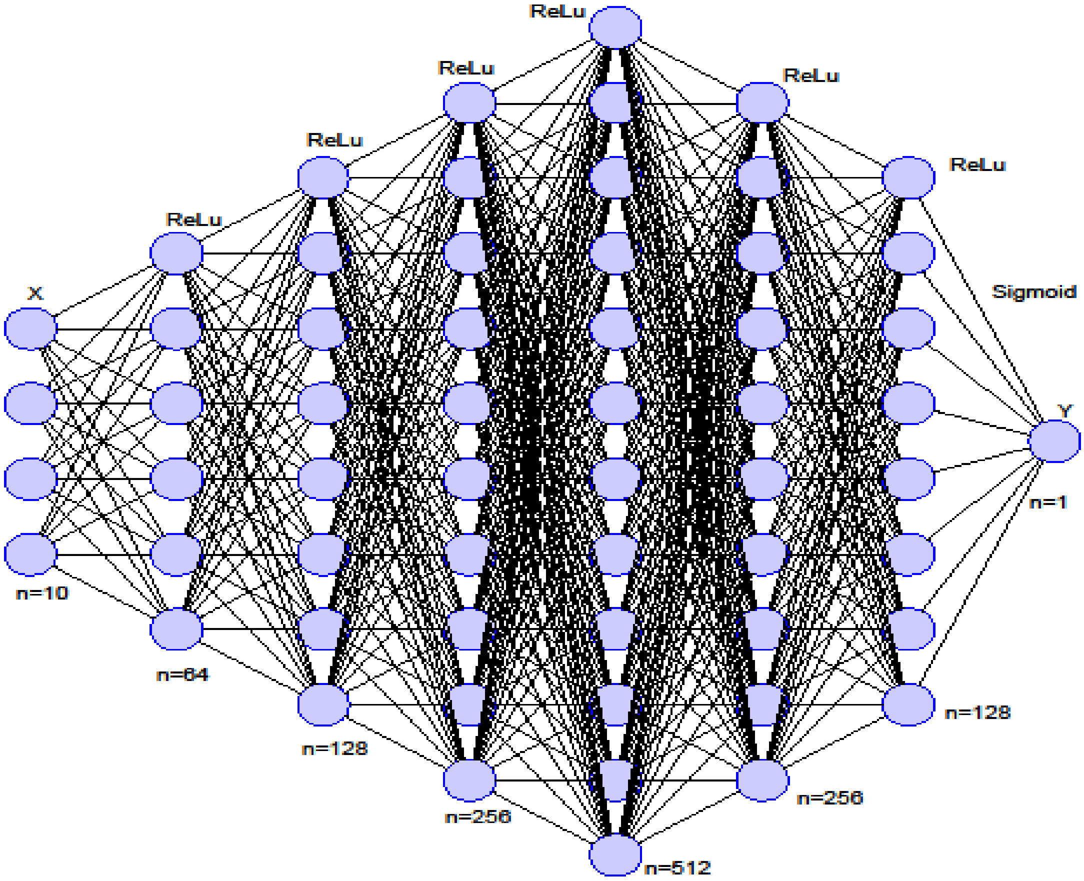


Fig 5: Layer diagram of CBFS-DNN

The performance of the CBFS-DNN is superior compared to other classifiers. By including a single neuron with sigmoid activation in the output layer while choosing "binary_crossentropy" as the loss function, the neural network in Figure 4 has the capability of carrying out binary classification. The significance of "binary_crossentropy" is that penalty becomes considerably bigger if the anticipated probability is actually wrong. The cross-entropy loss function exponentially increases the penalty for incorrect outputs during training to more aggressively push the weights and biases in desired direction. Number of hidden layers used are 64, 128, 256, 512, 256, 128 respectively. The neural network is trained for about 100 epochs.

Further it uses the Adam optimizer which is a crucial algorithm in deep learning, dedicated to enhancing the precision of neural networks by fine-tuning the model's learnable parameters. Its name, "Adam," signifies "Adaptive Moment Estimation," highlighting its ability to adapt the learning rate and estimate moments. This optimizer serves as an extension of widely-used stochastic gradient descent algorithm, which is responsible for adjusting weights in a neural network. In summary, the Adam optimizer is a potent tool for enhancing both the accuracy and efficiency of deep learning models.

Table 3: Hyperparameters of best performing CBFS-DNN

Parameters	**Values**
Activation function	'relu'
Loss	'binary_crossentropy'
Hidden layer sizes	(64,128,256,512,256,128)
Epochs	100
Optimizer	'Adam'

4. Results and Discussion

In this section, the research findings are discussed while also providing an in-depth analysis.

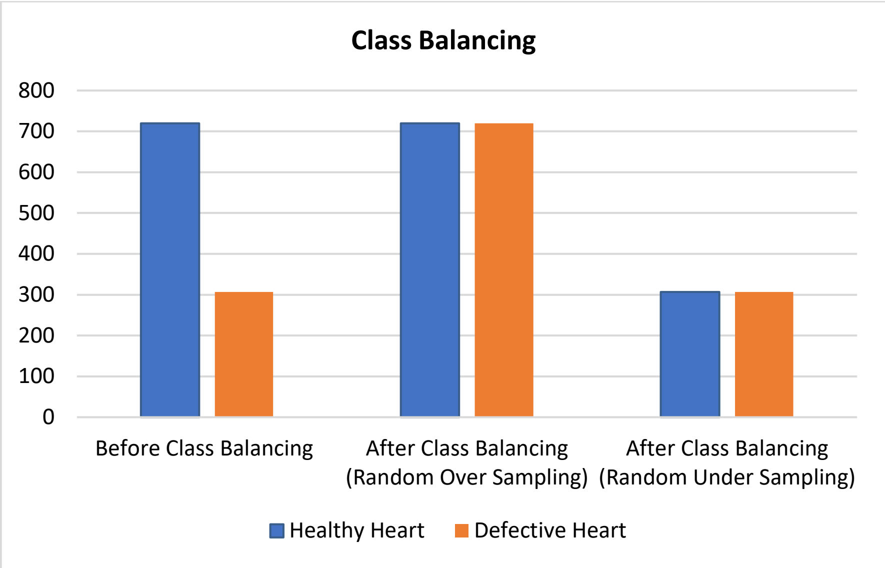


Fig 5: Class Balancing

Outlier handling is crucial to ensure accurate and reliable results in data analysis and modeling. It involves identifying and treating extreme values that can adversely impact statistical measures and machine learning algorithms.

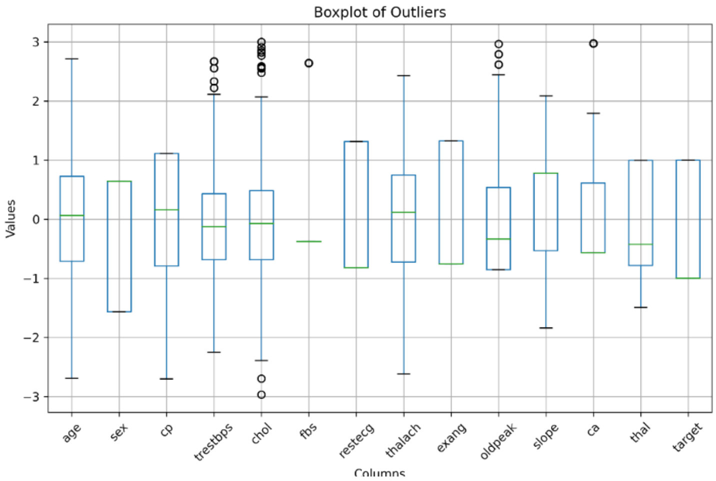
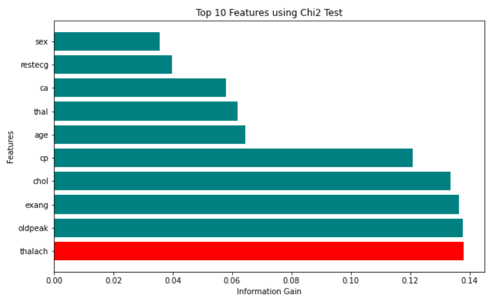


Fig 6: Outlier Handling

All parameters are studied from different datasets. It has been observed that most researchers have used either 13 or 14 features on a single dataset. We have concluded that the model works faster on the reduced number of features without compromising the efficiency of the model. This also helped in reducing the processing time which achieved the main objective of early detection of CVD. Imbalanced data, where the number of healthy individuals is much higher than CVD patients, can result in biased models that may neglect crucial patterns in the minority class. To avoid this class balancing is performed.



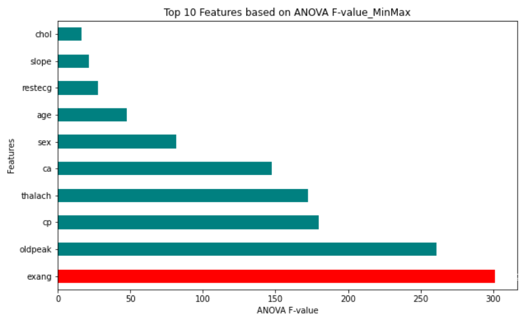


Fig 7(a) Top 10 Features using CHI2 Test

Fig 7(b) Top 10 Characteristics using ANOVA Test

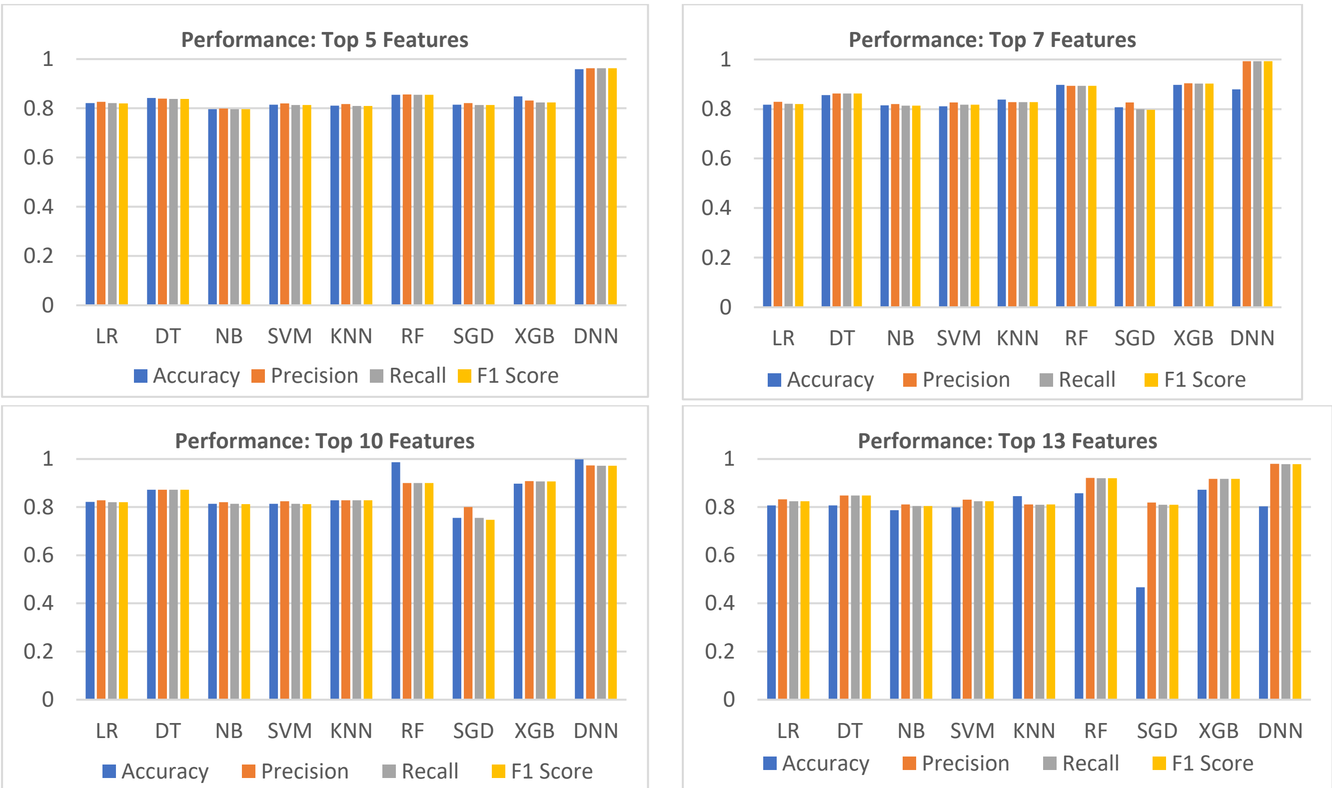


Fig 8: Performance comparison of Top-5, Top-7, Top-10 and all 13 features

Based on performance evaluation graph shown in Figure 7, system has utilized top 10 features for building the final model. By considering the 10 features the system have evaluated the model. And achieved an accuracy of 99.27%. System has also analyzed the other evaluation factors such as precision, recall and F1 score.

3.4. Evaluation parameters

By dividing the total number of accurate projections by the total number of predictions, as shown in Equation (1), one can determine reliability, which is expressed as a ratio.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

The proportion of True Positives to All Positives is how precision is measured, as demonstrated in Equation (2). In the context of our problem statement, this represents the proportion of correctly identified heart disease patients out of all the patients who genuinely have the condition.

$$Precision = \frac{TP}{TP + FP}$$

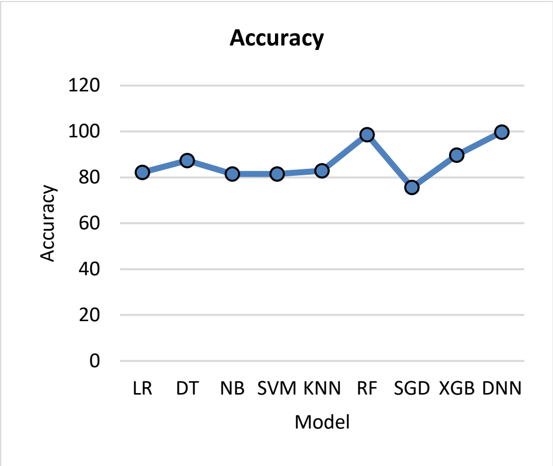
Recall, expressed as a measure, indicates our model's ability to correctly identify True Positives, as demonstrated in Equation (3). It essentially reveals how accurately our model identifies individuals with heart disease among those who genuinely have the condition.

$$Accuracy = \frac{TP}{TP + FN}$$

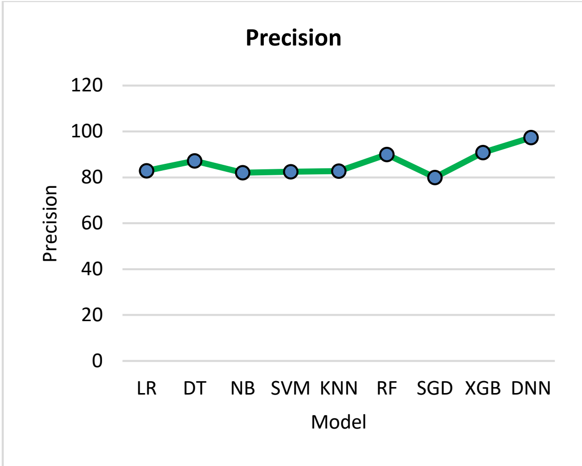
The F1 score, represented by Equation (4), functions as precise and recalling harmonic mean. It amalgamates both precision and recall into a unified metric, offering a well-balanced evaluation of a model's performance. This metric proves particularly valuable in situations where there is an uneven distribution between the positive and negative classes or when achieving a balance between precision and recall is a priority.

$$\text{F1 Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (3)$$

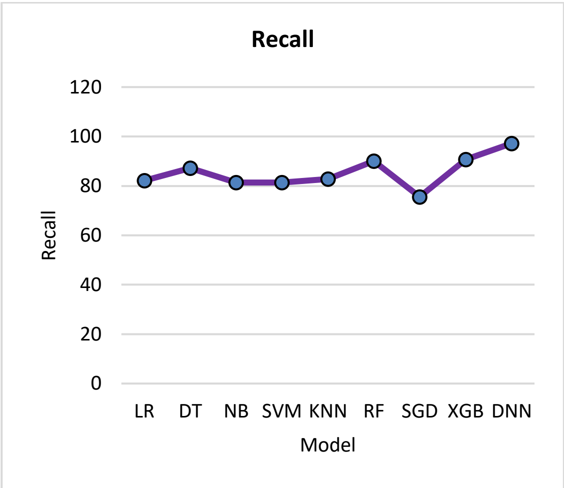
Figure 9, depicted below, illustrates the evaluation metrics for CBFS-DNN. As indicated in Figure 9(a), CBFS-DNN achieved the highest accuracy, reaching an impressive 99.79%. Similarly, the model's Precision, Recall, and F1 Score are showcased in Figure (b), Figure (c), and Figure (d), with values of 97.3%, 97.2%, and 97.2%, respectively.



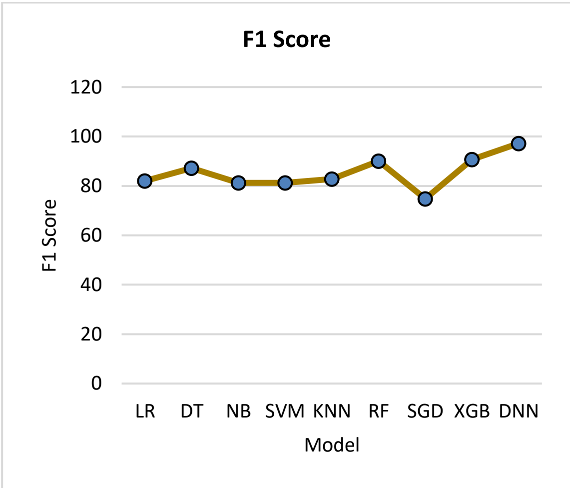
(a) Accuracy of all models



(b) Precision of all models



(c) Recall of all models



(d) F1 Score of all models

Fig 9: Evaluation Parameters (a) Accuracy, (b) Precision, (c) Recall and (d) F1 Score of all models

5. Conclusion

This study introduced a novel methodology called CBFS-DNN, which utilized a hybrid dataset for model training. The integration of ANOVA with machine learning proved to be effective in extracting important features. The proposed CBFS-DNN model displayed outstanding performance, achieving an impressive accuracy rate of 99.79% while utilizing only ten features and processing the data in a fraction of a second. Additionally, the model demonstrates Precision, Recall, and F1 Score values of 97.3%, 97.2%, and 97.2%, respectively. As the research progresses, the emphasis will be on further refining accuracy through real-time dataset testing. Overall, this research represents a significant advancement in the field of cardiovascular disease prevention by maintaining critical features within specified threshold values.

References

[1] E. Ahmad, "Cardiovascular Diseases (CVDs)

Detection using Machine Learning Algorithms," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 8, no. 6, pp. 2341–2346, 2020, doi: 10.22214/ijraset.2020.6376.

[2] D. Prabhakaran *et al.*, "The changing patterns of cardiovascular diseases and their risk factors in the states of India: the Global Burden of Disease Study 1990–2016," *Lancet Glob. Heal.*, vol. 6, no. 12, pp. e1339–e1351, Dec. 2018, doi: 10.1016/S2214-109X(18)30407-8.

[3] J. Poorolajal, "Neglected major causes of death much deadlier than COVID-19," *J. Res. Health Sci.*, vol. 20, no. 2, pp. 19–21, 2020, doi: 10.34172/jrhs.2020.12.

[4] J. M. Rippe, "AnAlytic," 2018, doi: 10.1177/1559827618812395.

[5] S. Frančula-Zaninović and I. A. Nola, "Send Orders for Reprints to reprints@benthamscience.ae Management of Measurable Variable Cardiovascular Disease' Risk Factors," *Curr. Cardiol. Rev.*, vol. 14, pp. 153–163, 2018, doi: 10.2174/1573403X14666180222102312.

[6] F. Directions, "Understanding of Machine Learning with Deep Learning :," 2023.

[7] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0217-0.

[8] I. Arel, D. Rose, and T. Karnowski, "Deep machine learning-A new frontier in artificial intelligence research," *IEEE Comput. Intell. Mag.*, vol. 5, no. 4, pp. 13–18, 2010, doi: 10.1109/MCI.2010.938364.

[9] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," *Sensors*, vol. 21, no. 11, pp. 1–55, 2021, doi: 10.3390/s21113758.

[10] J. Huang, J. Chai, and S. Cho, "Deep learning in finance and banking: A literature review and classification," *Front. Bus. Res. China*, vol. 14, no. 1, 2020, doi: 10.1186/s11782-020-00082-6.

[11] C. Gambella, B. Ghaddar, and J. Naoum-Sawaya, "Optimization problems for machine learning: A survey," *Eur. J. Oper. Res.*, vol. 290, no. 3, pp. 807–828, 2021, doi: 10.1016/j.ejor.2020.08.045.

[12] Q. Vuong, "Machine Learning for Robotic Manipulation," 2021, [Online]. Available: http://arxiv.org/abs/2101.00755

[13] F.-G. Yuan, S. A. Zargar, Q. Chen, and S. Wang, "Machine learning for structural health monitoring: challenges and opportunities," vol. 1137903, no. April 2020, p. 2, 2020, doi: 10.1117/12.2561610.

[14] M. Peng *et al.*, "Prediction of cardiovascular disease risk based on major contributing features," *Sci. Rep.*, vol. 13, no. 1, pp. 1–11, 2023, doi: 10.1038/s41598-023-31870-8.

[15] A. Javeed, S. S. Rizvi, S. Zhou, R. Riaz, S. U. Khan, and S. J. Kwon, "Heart risk failure prediction using a novel feature selection method for feature refinement and neural network for classification," *Mob. Inf. Syst.*, vol. 2020, 2020, doi: 10.1155/2020/8843115.

[16] P. Rahman, A. Rifat, I. A. Chy, M. M. Khan, M. Masud, and S. Aljahdali, "Machine Learning and Artificial Neural Network for Predicting Heart Failure Risk," *Comput. Syst. Sci. Eng.*, vol. 44, no. 1, pp. 757–775, 2022, doi: 10.32604/csse.2023.021469.

[17] A. M. Qadri, A. Raza, K. Munir, and M. S. Almutairi, "Effective Feature Engineering Technique for Heart Disease Prediction With Machine Learning," *IEEE Access*, vol. 11, no. May, pp. 56214–56224, 2023, doi: 10.1109/ACCESS.2023.3281484.

[18] B. P. Doppala, D. Bhattacharyya, M. Chakkravarthy, and T. Kim, "A hybrid machine learning approach to identify coronary diseases using feature selection mechanism on heart disease dataset," *Distrib. Parallel Databases*, vol. 41, no. 1, pp. 1–20, 2023, doi: 10.1007/s10619-021-07329-y.

[19] L. A. Alqahtani, H. M. Alotaibi, I. U. Khan, and N. Aslam, "Automated prediction of Heart disease using optimized machine learning techniques," *2020 11th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2020*, no. October 2020, pp. 0298–0302, 2020, doi: 10.1109/UEMCON51285.2020.9298051.

[20] G. Magesh and P. Swarnalatha, "Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction," *Evolutionary Intelligence*, vol. 14, no. 2. pp. 583–593, 2021. doi: 10.1007/s12065-019-00336-0.

[21] P. Rani, R. Kumar, A. Jain, and R. Lamba, "Taxonomy of Machine Learning Algorithms and Its Applications," *J. Comput. Theor. Nanosci.*, vol. 17, no. 6, pp. 2508–2513, Sep. 2020, doi: 10.1166/JCTN.2020.8922.

[22] S. P. Patro, G. S. Nayak, and N. Padhy, "Heart disease prediction by using novel optimization algorithm: A supervised learning prospective," *Informatics Med. Unlocked*, vol. 26, 2021, doi: 10.1016/j.imu.2021.100696.

[23] M. S. Nawaz, B. Shoaib, and M. A. Ashraf, "Intelligent Cardiovascular Disease Prediction Empowered with Gradient Descent Optimization," *Heliyon*, vol. 7, no. 5, p. e06948, 2021, doi: 10.1016/j.heliyon.2021.e06948.

[24] D. P. Yadav, P. Saini, and P. Mittal, "Feature Optimization Based Heart Disease Prediction using Machine Learning," *2021 5th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2021*, no. March, 2021, doi: 10.1109/ISCON52037.2021.9702410.

[25] H. B. Kibria and A. Matin, "The severity prediction of the binary and multi-class cardiovascular disease

− A machine learning-based fusion approach,” *Comput. Biol. Chem.*, vol. 98, 2022, doi: 10.1016/j.compbiolchem.2022.107672.

[26] V. V. R. Karna, S. Paramasivam, I. Elamvazuthi, H. N. Chua, A. A. Aziz, and P. Satyamurthy, “Novel Feature Engineering for Heart Disease Risk Prediction Using Optimized Machine Learning,” 2022 Int. Conf. Futur. Trends Smart Communities, ICFTSC 2022, pp. 158–163, 2022, doi: 10.1109/ICFTSC57269.2022.10040063.

[27] Mwangi, J., Cohen, D., Costa, R., Min-ji, K., & Suzuki, H. Optimizing Neural Network Architecture for Time Series Forecasting. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/132

[28] Barwal, R. K. ., Raheja, N. ., Bhiyana, M. ., & Rani, D. . (2023). Machine Learning-Based Hybrid Recommendation (SVOF-KNN) Model For Breast Cancer Coimbra Dataset Diagnosis. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1s), 23–42. https://doi.org/10.17762/ijritcc.v11i1s.5991

Biographies Of Authors

	**Ms Shivganga Udhan** is an Assistant Professor in Dr.D.Y.Patil Institute of Engineering Management and Research Akurdi, Pune (India). She has done her ME in Computer Network from Sihngad College of Engineering, Savitribai Phule Pune University (India). She has more than 10 years of teaching experience. She has the life membership of Indian Society for Technical Education (ISTE). She also has the membership of International Association of Engineers (IAENG). Her research interests are inNeural Network, Arti_cial Intelligence and Machine Learning.
	**Dr.Bankat M.Patil Professor** Department of Computer Science Engineering, M.B.E. Society's College of Engineering, Ambajgai. (India). He has completed his Ph.D. in Computer Science (Data Mining) from Indian Institute of Technology, Roorkee. He has done his M. Tech in Software Engineering from Mysore University. He has wide experience of more than 20 years. He has the life membership of Indian Society for Technical Education (ISTE). He is also member of IEEE Society's the world's leading professional association for the advancement oftechnology.(IEEE). His research interest are Parallel processing, Data Mining, Soft Computing, Decision Support Systemin Medicine, Computer networking, Cloud Computing