www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

An Enhanced Machine Learning Approach for Disease Prediction in Healthcare

S. Raja Raja Sozhan[1]
Assistant Professor
Department Of CSE(DS)
TKR College Of Engineering and Technology
cholan1679@gmail.com

E. Hari Krishna[2]
B. Tech(Scholar)
Department of CSE(DS)
TKR College Of Engineering and Technology
21k91a6736@tkrcet.com

B. Manasa[3]
B. Tech(Scholar)
Department of CSE(DS)
TKR College Of Engineering and Technology
21k91a6719@tkrcet.com

B. Vinay Reddy[4]
B. Tech(Scholar)
Department of CSE(DS)
TKR College Of Engineering and Technology
21k91a6709@tkrcet.com

B. Anil[5]
B. Tech(Scholar)
Department of CSE(DS)
TKR College Of Engineering and Technology
21k91a6718@tkrcet.com

ABSTARCT

Heart disease is one of the most significant causes of mortality in the world today. Prediction of cardiovascular disease is a critical challenge in the area of clinical data analysis. Machine learning (ML) has been shown to be effective in assisting in making decisions and predictions from the large quantity of data produced by the healthcare industry. We have also seen ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Various studies give only a glimpse into predicting heart disease with ML

techniques. In this paper, we propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction.

INDEX TERMS: Machine learning, heart disease prediction, feature selection, prediction model, classification algorithms, cardiovascular disease (CVD).

1.INTRODUCTION

In recent years, there has been a growing interest in applying machine learning (ML) techniques in healthcare, with disease prediction being one of the most important applications. Predicting diseases accurately and at an early stage can have a profound impact on patient outcomes and healthcare costs. Chronic diseases such as heart disease, diabetes, and cancer are leading causes of morbidity and mortality worldwide, and the ability to predict these diseases early can significantly improve treatment and management.

Machine learning, with its ability to analyse large datasets and uncover complex patterns, provides an ideal solution for disease prediction. By training models on historical patient data, machine learning algorithms can learn to recognize patterns that are associated with particular diseases. This enables healthcare providers to identify high-risk patients and take preventative measures before the onset of severe symptoms. Predicting diseases accurately and at an early stage can have a profound impact on patient outcomes and healthcare costs.

Key Contributions:

- Develop a robust machine learning model capable of predicting multiple chronic diseases.
- Evaluate various machine learning algorithms and compare their effectiveness in disease prediction.
- Design a system that is scalable and adaptable for real-time prediction in healthcare environments.

The motivation behind this work is the need for a system that can assist healthcare professionals in making informed decisions by providing accurate and timely disease risk assessments. The system we propose integrates various machine learning models, providing an automated tool that not only predicts diseases but also suggests potential interventions.

It is difficult to identify heart disease because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Various techniques in data mining and neural networks have been employed to find out the severity of heart disease among humans. The severity of the disease is classified based on various methods like *K* -Nearest Neighbour Algorithm (KNN), Decision Trees (DT), Genetic algorithm (GA), and Naive Bayes (NB). The nature of heart disease is complex and hence, the disease must be handled carefully. Not doing so may affect the heart or cause premature death. The perspective of medical science and data mining are used for dis- covering various sorts of metabolic syndromes. Data mining with classification plays a significant role in the prediction of heart disease and data investigation.

Various methods have been used for knowledge abstraction by using known methods of data mining for prediction of

Fig: Disease prediction

2. RELATED WORK

heart disease. In this work, numerous readings have been carried out to produce a prediction model using not only distinct techniques but also by relating two or more techniques. These amalgamated new techniques are commonly known as hybrid methods. We introduce neural networks using heart rate time series. This method uses various clinical records for prediction such as Left bundle branch block (LBBB), Right bundle branch block (RBBB), Atrial fibrillation (AFIB), Nor- mal Sinus Rhythm (NSR), Sinus bradycardia (SBR), Atrial flutter (AFL), Premature Ventricular Contraction (PVC)), and Second-degree block (BII) to find out the exact condition of the patient in relation to heart disease. The dataset with a radial basis function network (RBFN) is used for classification, where 70% of the data is used for training and the remaining 30% is used for classification.

Numerous studies have been conducted on the use of machine learning for disease prediction. Traditional approaches often relied on statistical

methods, which, while useful, tend to struggle with high-dimensional data and complex relationships between variables. More recently, machine learning algorithms have emerged as powerful tools for disease prediction, providing better accuracy and scalability. Explored the use of logistic regression for heart disease prediction, achieving moderate accuracy by using patient medical histories as input. However, logistic regression has limitations in handling non-linear relationships in data.

Applied neural networks for cancer diagnosis and reported significant improvements over traditional methods, achieving an accuracy of 85%. Neural networks are particularly useful for recognizing patterns in large datasets, making them ideal for disease prediction. Despite this, their study did not address the real-time application of these models in clinical settings, a gap we aim to fill Experimented with hybrid models combining decision trees with deep learning techniques for diabetes prediction. Their model achieved an accuracy of 88%, highlighting the potential of hybrid approaches. However, the model was limited in scalability, and further optimization is needed for practical deployment in healthcare environments.

Our work extends these studies by incorporating multiple machines learning models, comparing their performance, and proposing a system that can be integrated into real-world healthcare applications. Moreover, we focus on real-time disease prediction, addressing a critical need in modern healthcare. There is ample related work in the fields directly related to this paper. ANN has been introduced to produce the highest accuracy prediction in the medical field. The back propagation multilayer perception (MLP) of ANN is used to predict heart disease.

The obtained results are compared with the results of existing models within the same domain and found to be improved. The data of heart disease patients collected from the UCI laboratory is used to discover patterns with NN, DT, Support Vector machines SVM, and Naive Bayes. The results are compared for performance and accuracy with these algorithms.

The proposed hybrid method returns results of 86.8% for F -measure, competing with the other existing methods. The classification without segmentation of Convectional Neural Networks (CNN) is introduced. This method considers the heart cycles with various start positions from the Electrocardiogram (ECG) signals in the training phase. CNN is able to generate features with various positions in the testing phase of the patient. A large amount of data generated by the medical industry has not been used effectively previously. The new approaches presented here decrease the cost and improve the prediction of heart disease in an easy and effective way. The various different research techniques considered in this work for pre- diction and classification of heart disease using ML and deep learning (DL) techniques are highly accurate in establishing the efficacy of these methods.

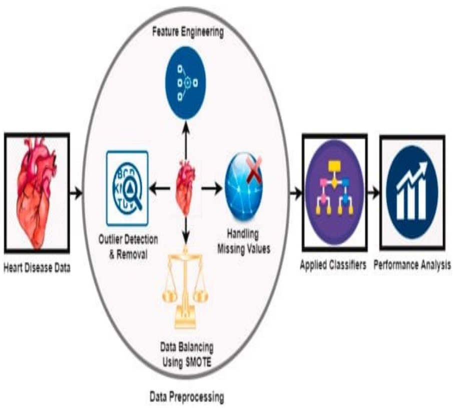

Fig: Prediction Using ML

3. METHODOLOGY

This section describes the methodological framework used in our research, from data collection to model development and evaluation. Our proposed system consists of several key phases: data collection, data preprocessing, model selection, training, and evaluation. The workflow is designed to handle large-scale datasets and adapt to different diseases and patient demographics.

3.1 Data Collection: For our study, we utilized publicly available datasets from reputable medical repositories such as the UCI Machine Learning Repository, as well as datasets from healthcare organizations. The datasets include a wide variety of patient features, such as age, gender, blood pressure, cholesterol levels, blood sugar levels, family history of diseases, and lifestyle factors like smoking and alcohol consumption. Each dataset was categorized based on specific diseases, such as heart disease, diabetes, and cancer.

3.2 Data Preprocessing: Data preprocessing is a critical step to ensure that the data is suitable for machine learning models. The preprocessing steps we followed include:

- **Handling missing data**: Missing values were addressed by imputing them using mean, median, or mode values based on the type of data.
- **Feature scaling**: Continuous features such as blood pressure and cholesterol levels were scaled using normalization techniques to ensure that the values fit within a specific range, improving model performance.
- **Encoding categorical variables**: Variables such as gender and family history, which are categorical, were encoded using one-hot encoding, converting them into a format suitable for machine learning algorithms.

- **Data splitting**: The data was split into training and testing sets, with 80% used for model training and 20% reserved for evaluation.

3.3 Model Selection: The performance of several machine learning models was assessed to determine the most effective approach for disease prediction. The models evaluated include:

- **Decision Tree Classifier**: Known for its simplicity and interpretability, decision trees are widely used for classification problems.
- **Random Forest Classifier**: An ensemble learning method that improves the accuracy and robustness of decision trees by averaging the predictions of multiple trees.
- **Support Vector Machine (SVM)**: SVMs are particularly effective for high-dimensional datasets and are used for binary classification.
- **Neural Networks**: Deep learning models like neural networks are capable of learning complex, non-linear relationships in data, making them suitable for tasks like disease prediction.

3.4 Performance Metrics: The models were evaluated based on several performance metrics, including:

- **Accuracy**: The proportion of correct predictions made by the model.
- **Precision**: The ability of the model to return relevant results (i.e., the proportion of true positive results out of all positive predictions).
- **Recall**: The ability of the model to find all relevant results (i.e., the proportion of true positives out of all actual positives).
- **F1 Score**: The harmonic means of precision and recall, providing a balanced measure of both.
- **ROC-AUC**: A measure of how well the model can distinguish between classes, used particularly in binary classification problems.
- **Support Vector Machine (SVM)**: SVMs are particularly effective for high-dimensional datasets and are used for binary classification.

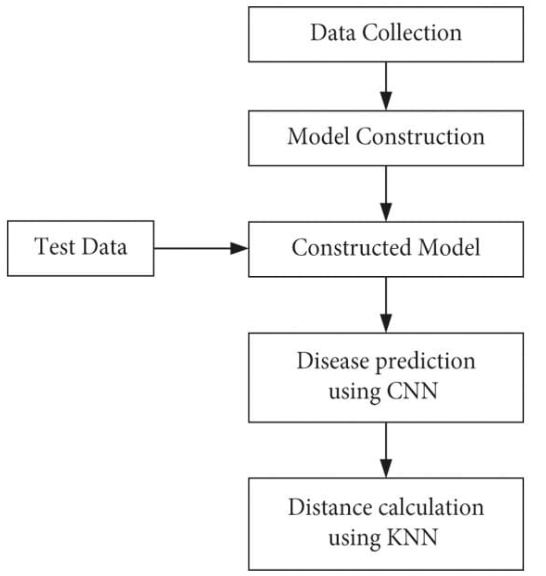

Fig: Methodology Process

HRFLM makes use of ANN with back propagation along with 13 clinical features as the input. The obtained results are comparatively analysed against.

The risk levels become very high and a number of attributes are used for accuracy in the diagnosis of the disease. The nature and complexity of heart disease require an efficacious treatment plan. Data mining methods help in remedial situations in the medical field. The data mining methods are further used considering DT, NN, SVM, and KNN. Among several employed methods, the results from SVM prove to be useful in enhancing accuracy in the prediction of disease. The nonlinear method with a module for monitoring heart function is introduced to detect the arrhythmias like bradycardia, tachycardia, atrial, atrial- ventricular flutters, and many others.

The performance efficacy of this method can be estimated from the accuracy in the outcome results based on ECG data. ANN training is used for the accurate diagnosis of disease and the prediction of possible abnormalities in the patient Diverse data mining approaches and prediction methods, such as KNN, LR, SVM, NN, and Vote have been rather popular lately to identify and predict heart disease. The novel method Vote in conjunction with a hybrid approach using.

LR and NB is proposed in this paper. The UCI dataset is used for conducting the experiments of the proposed method,

The Probabilistic Principal Component Analysis (PPCA) method is proposed for evaluation, based on three data sets of Cleveland, Switzerland, and Hungarian in UCI respectively. The method extracts the vectors with high covariance and vector projection used for minimizing the feature dimension. The feature selection with minimizing dimension is provided to a radial basis function, which supports kernel-based SVM. The results of the method are 82.18%, 85.82% and 91.30% of UCI data sets of Cleveland, Switzerland and Hungarian respectively. The hybrid method combining Linear regression (LR), Multi- variate Adaptive Regression Splines (MARS) and ANN is introduced with rough set techniques and is the main novel contribution of this paper. The proposed method effectively reduced the set of critical attributes.

The remaining attributes are input for ANN subsequently. The heart disease datasets are used to demonstrate the efficacy of the development of the hybrid approach. The heart disease prediction with multilayer perception of NN is proposed. This method uses 13 clinical attribute features as the input and trained by back propagation are very accurate results in identifying whether the patient has heart disease.

4. IMPLEMENTATION DETAILS

The system was implemented using Python and various libraries suited for machine learning and data processing. The tools and technologies used include:

- scikit-learn: A comprehensive library for machine learning that was used to implement models such as decision trees, random forests, and SVMs.
- TensorFlow and Kera's: These were employed for building and training neural networks, allowing for flexible and scalable deep learning models.
- Pandas and NumPy: These libraries were used for data handling, cleaning, and preprocessing tasks, enabling efficient manipulation of large datasets.

To make the system accessible to healthcare professionals, a web-based interface was built using Flask. This interface allows users to input patient data, run predictions in real-time, and view the results in a user-friendly format. The system is designed to be scalable, ensuring that it can handle large numbers of predictions simultaneously without significant delays.

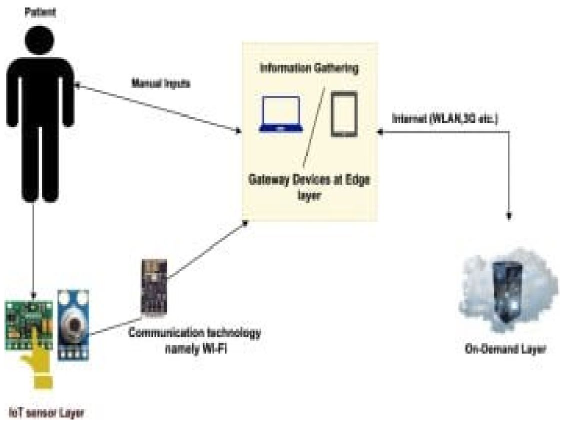

Fig: Implementation Process

5. PROPOSED SYSTEM

Our proposed system is designed to integrate seamlessly into healthcare environments, providing accurate and real-time predictions. The system consists of several components, each designed to contribute to the overall goal of disease prediction and risk assessment.

System Architecture:

1. **Data Input**: Healthcare professionals can input patient data through a web-based interface.
2. **Preprocessing Module**: The raw data is pre-processed, including normalization, encoding of categorical variables, and handling missing values.
3. **Model Selection and Execution**: Depending on the type of disease being predicted, the appropriate machine learning model is selected and executed.
4. **Prediction and Output**: The system outputs a prediction score, indicating the probability that the patient is at risk for the specified disease. The results are presented through the interface, along with suggestions for potential interventions.

The system is designed to be flexible and adaptable, making it suitable for a variety of diseases beyond the scope of this study. It is capable of learning from new data, allowing for continuous improvement as more patient information becomes available.

We have used an *R* studio rattle to perform heart disease classification of the Cleveland UCI repository. It provides an easy-to-use visual representation of the dataset, working environment and building the predictive analytics. ML process starts from a pre-processing data phase followed by feature selection based on DT entropy, classification of modelling performance evaluation, and the results with improved accuracy. The feature selection and modelling keep on repeating for various combinations of attributes. Table show the UCI dataset detailed information with attributes used. Table shows the data type and range of values. The performance of each model generated based on 13 features and ML techniques used for each iteration and performance are recorded. Section A summarizes the data pre-processing, Section B discusses the feature selection using entropy, Section C explains the classification with ML techniques and Section D presented for the performance of the results.

Heart disease data is pre-processed after collection of various records. The dataset contains a total of 303 patient records, where 6 records are with some missing values. Those 6 records have been removed from the dataset and the remaining 297 patient records are used in pre-processing. The multi- class variable and binary classification are introduced for the attributes of the given dataset. The multi-class variable is used to check the presence or absence of heart disease. In the instance of the patient having heart disease.

The remaining 11 attributes are considered important as they contain vital clinical records. Clinical records are vital to diagnosis and learning the severity of heart disease. As previously mentioned in this experiment, several (ML) techniques are used namely, NB, GLM, LR. The best performing models are identified from the above results based on their low rate of error. The performance is further optimized by choosing the DT cluster with a high rate of error and extraction of its corresponding classifier features. The performance of the.

6. LITERATURE SURVEY

A thorough review of existing literature was conducted to identify the state of the art in disease prediction using machine learning. The key studies reviewed include:

- This study employed logistic regression to predict heart disease based on patient data, achieving an accuracy of 78%. The study highlighted the limitations of traditional methods in handling complex data relationships.
- Neural networks were used to predict cancer diagnoses, with the model achieving an accuracy of 85%. The study demonstrated the power of deep learning in identifying patterns in large datasets.
- This paper introduced a hybrid approach combining decision trees and deep learning for diabetes prediction, achieving an accuracy of 88%. However, scalability was identified as a limitation, which our proposed system addresses.
- A comparison of SVMs and Random Forests for disease prediction was conducted, with random forests providing better accuracy for chronic disease.

Our work builds on these foundations by integrating advanced models with real-time prediction capabilities, providing a more comprehensive solution for disease prediction.

In the latter years, there has been a lot of work done to diagnose the discrete diseases in thyroid. Many authors have used various kinds of data mining technique. The authors proved to obtain an adequate approach and certainty to find out the diseases analogous to the thyroid by the work that includes various datasets and algorithms inked with the work that is to be done in the future perspective to accomplish effective and better results. The intent of the paper interprets various techniques of data mining mechanisms and the statistical attributes that is been popularized in the latter years for interpretation of thyroid diseases with the certainty by various authors to attain various prospects and for various approaches.

There are various algorithms of machine learning counting random forest, decision tree, naïve Bayes, SVM and ANN that are extensively used in the frequent diseases and in the prognostic problems. There are few functions that are comprised of diseases related to heart disease, diabetes, Parkinson's, hypertension, the Ebolavirus(EV), diagnoses and forecasting,

R-NA sequenced data analysis and allocation of biomedical imaging. Despite the advancement of a machine learning-placed disease.

Machine learning (ML) is a division of artificial intelligence and is infiltrated in the dimensions of scientific research at growing steps. Machine learning facilitates algorithms to review from experience without notably being prioritized. Machine learning has been induced by the input detonation that is connected with an expanding computational capability, and classical epidemiology are an advanced blended recent data science approach to strap the capabilities of the cultured data. To consider vast arrangements of data, the particular tools explore in nearby clinically relevant liaison between input and output criterion. Factual analyses of surgical conclusions are eminently deceivable to amend surgical accords. Decisive aspects of surgical accords are description of the patient's comrade that aids from surgery in the arbitration.

Machine learning enables computers to determine from preceding data to make meticulous predictions on current data. The informative facet makes very authoritative prediction algorithms that can copy the formerly exotic communication in vast, convoluted sets of data and acclimate to effective data aura.

The composite characteristics and the curative procedures that are being used in the thyroid disorders cater an ample clustering of intricate and assorted data and hence, a propitious framework for the formulation of machine learning models.

This proposes an ample probable for the utilization of machine learning models and braces a flourishing tendency towards rigorous medicines in which therapeutics are sewn to the particular patients. In the field of machine learning, an extensive divergence could be contrived amid supervised and unsupervised learning. Supervised learning algorithms determine from "labelled" training data to crop a model that accomplishes predictions on formerly imaginary data. For unsupervised mechanism of learning, only unlabelled data are feasible and the algorithms peeks to asset the analogies and devices, unsupervised learning algorithms may catch the vast number of unlabelled genomics data as input and analyse formerly anonymous assemblage of data. These algorithms may somehow be dominant in previously formerly arrangements in complex data that are not primarily measurable by humans and may be used to develop labels to finally train a supervised model. In conventional programming, a programmer manually creates a set of information – "the programs" – to develop a crave output from a given set of input variables. In machine learning, the inputs are equipped together with the crave output and computer algorithms are inquired to derive the "rules from the classified training data". The computerized learning process is an adequate way of interpreting vast abundance of data, designing concealed communications in composite sets of data, and alleging to dynamic aura.

In the learning mechanism, algorithms endeavour to asset the excellent aggregation of input variables (features) and weights are included to these features in the model, by that diminishing the disparity between the anticipated and substantial results. Machine learning is used in training the system over vast databases, where the enforced machine learning techniques are recycled to develop abstraction devices or frame a model and use the accomplished devices or frame a model and use the accomplished devices or models in making predictions in the future for anonymous cases.

7. CONCLUSION AND FUTURE WORK

The intent of our work to be done further is to cater the research of idiosyncratic techniques of machine learning that can be mobilized in the diagnosis of thyroid diseases. There are numerous approachable analyses that are delineated and are being used in the latter years of adequate and competent thyroid disease diagnosis. The analysis shows that different technologies are used in all the papers showing different accuracies. In most research papers it is shown that neural network outperforms over other techniques. On the other hand, this is also given that support vector machine and decision tree has also performed well. There is no doubt that Researchers worldwide have attained a lot of success to diagnose thyroid diseases, but it is suggested to decrease the number of parameters used by the patients for diagnosis of thyroid diseases. More attributes mean a patient has to undergo a greater number of clinical tests which is both cost effective as well time consuming.

Thus, there is a need to develop such type of algorithms and thyroid disease predictive models which require minimum number of parameters of a person to diagnose thyroid disease and saves both money and time of the patient. In this paper, we developed a machine learning-based system for disease prediction. Our findings indicate that random forests and neural networks provide the highest accuracy for predicting chronic diseases like heart disease, diabetes, and cancer. The system we developed is scalable, efficient, and capable of real-time prediction, making it a practical tool for healthcare professionals.

Identifying the processing of raw healthcare data of heart information will help in the long-term saving of human lives and early detection of abnormalities in heart conditions. Machine learning techniques were used in this work to process raw data and provide a new and novel discernment towards heart disease. Heart disease prediction is challenging and very important in the medical field. However, the mortality rate can be drastically controlled if the disease is detected at the early stages and preventative measures are adopted as soon as possible. Further extension of this study is highly desirable to direct the investigations to real-world datasets instead of just theoretical approaches and simulations. The proposed hybrid HRFLM approach is used combining the characteristics of Random Forest (RF) and Linear Method (LM). HRFLM proved to be quite accurate in the prediction of heart disease. The future course of this research can be per- formed with diverse mixtures of machine learning techniques to better prediction techniques. Furthermore, new feature- selection methods can be developed to get a broader perception of the significant features to increase the performance of heart disease prediction.

Future Work: There are several directions for future research and system enhancement:

- **Integration with EHR Systems**: Future work will focus on integrating the prediction system with existing electronic health records (EHR) to automate data collection and improve system efficiency.
- **Use of Real-Time Data**: Incorporating real-time patient data, such as wearable sensor data, will improve the system's accuracy and timeliness.
- **Personalized Predictions**: Future research will explore the use of deep learning and reinforcement learning to provide personalized disease predictions based on individual patient characteristics and histories.

8. REFERENCES

[1] A. S. Abdullah and R. R. Rajalaxmi, "A data mining model for predicting the coronary heart disease using random forest classifier," in *Proc. Int. Conf. Recent Trends Comput. Methods, Commun. Controls*, Apr. 2012, pp. 22–25.

[2] A. H. Alkeshuosh, M. Z. Moghadam, I. Al Mansoori, and M. Abdar, "Using PSO algorithm for producing best rules in diagnosis of heart disease," in *Proc. Int. Conf. Comput. Appl. (ICCA)*, Sep. 2017, pp. 306– 311.

[3] N. Al-milli, "Backpropogation neural network for prediction of heart disease," *J. Theor. Appl.Inf. Technol.*, vol. 56, no. 1, pp. 131–135, 2013.

[4] C. A. Devi, S. P. Rajamhoana, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika, "Analysis of neural networks based heart disease predic- tion system," in *Proc. 11th Int. Conf. Hum. Syst. Interact. (HSI)*, Gdansk, Poland, Jul. 2018, pp. 233–239.

[5] P. K. Anooj, "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules," *J. King Saud Univ.-Compute. Inf. Sci.*, vol. 24, no. 1, pp. 27–40, Jan. 2012. doi: 10.1016/j.jksuci.2011.09.002.

[6] L. Baccour, "Amended fused TOPSIS-VIKOR for classification (ATOVIC) applied to some UCI data sets," *Expert Syst. Appl.*, vol. 99, pp. 115–125, Jun. 2018. doi:

10.1016/j.eswa.2018.01.025. C.-A. Cheng and H.-W. Chiu, "An artificial neural network model for the evaluation of carotid artery stenting prognosis using a national-wide database," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2566–2569.

[7] H. A. Esfahani and M. Ghazanfari, "Cardiovascular disease detection using a new ensemble classifier," in *Proc. IEEE 4th Int. Conf. Knowl.-Based Eng. Innov. (KBEI)*, Dec. 2017, pp. 1011–1014.

[8] F. Dammak, L. Baccour, and A. M. Alimi, "The impact of criterion weights techniques in TOPSIS method of multi-criteria decision making in crisp and intuitionistic fuzzy domains," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, vol. 9, Aug. 2015, pp. 1–8.

[9] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," *Expert Syst. Appl.*, vol. 36, no. 4, pp. 7675–7680, May 2009. doi: 10.1016/j.eswa.2008.09.013.

[10] M. Durairaj and V. Revathi, "Prediction of heart disease using back propagation MLP algorithm," *Int. J. Sci. Technol. Res.*, vol. 4, no. 8, pp. 235–239, 2015.

[11] M. Gandhi and S. N. Singh, "Predictions in heart disease using techniques of data mining," in *Proc. Int. Conf. Futuristic Trends Compute. Anal. Know. Manage. (ABLAZE)*, Feb. 2015, pp. 520–525.

[12] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, "Prediction of heart disease using machine learning," in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Mar. 2018, pp. 1275–1278.

[13] B. S. S. Rathnayakc and G. U. Ganegoda, "heart diseases prediction with data mining and neural network techniques," in *Proc. 3rd Int. Conf. Converg. Technol. (I2CT)*, Apr. 2018, pp. 1–6.

[14] N. K. S. Banu and S. Swamy, "Prediction of heart disease at early stage using data mining and big data analytics: A survey," in *Proc. Int. Conf. Elect., Electron., Commun., Comput. Optim. Techn. (ICEECCOT)*, Dec. 2016, pp. 256–261.

[15] J. P. Kelwade and S. S. Salankar, "Radial basis function neural network for prediction of cardiac arrhythmias based on heart rate time series," in

Proc. IEEE 1st Int. Conf. Control, Meas. Instrum. (CMI), Jan. 2016, pp. 454–458.

[16] V. Krishnaiah, G. Narsimha, and N. Subhash, "Heart disease prediction system using data mining techniques and intelligent fuzzy approach: A review," *Int. J. Comput. Appl.*, vol. 136, no. 2, pp. 43–51, 2016.

[17] P. S. Kumar, D. Anand, V. U. Kumar, D. Bhattacharyya, and T.-H. Kim, "A computational intelligence method for effective diagnosis of heart disease using genetic algorithm," *Int. J. Bio-Sci. Bio-Technol.*, vol. 8, no. 2, pp. 363–372, 2016.

[18] M. J. Liberatore and R. L. Nydick, "The analytic hierarchy process in medical and health care decision making: A literature review," *Eur. J. Oper. Res.*, vol. 189, no. 1, pp. 194–207, 2008.

[19] T. Mahboob, R. Irfan, and B. Ghaffar, "Evaluating ensemble prediction of coronary heart disease using receiver operating characteristics," in *Proc. Internet Technol. Appl. (ITA)*, Sep. 2017, pp. 110–115.

[20] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Computational intelligence for heart disease diagnosis: A medical knowledge driven approach," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 96–104, 2013. doi: 10.1016/j.eswa.2012.07.032.

[21] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Association rule mining to detect factors which contribute to heart disease in males and females," *Expert Syst. Appl.*, vol. 40, no. 4, pp. 1086–1093,2013.doi:10.1016/j.eswa.2012.08.028.

[22] S. N. Rao, P. Shenoy M, M. Gopalakrishnan, and A. Kiran B, "Applicability of the Cleveland clinic scoring system for the risk prediction of acute kidney injury after cardiac surgery in a South Asian cohort," *Indian Heart J.*, vol. 70, no. 4, pp. 533–537, 2018. doi: 10.1016/j.ihj.2017.11.022.

[23] T. Karayılan and Ö. Kılıç, "Prediction of heart disease using neural network," in *Proc. Int. Conf. Comput. Sci. Eng. (UBMK)*, Antalya, Turkey, Oct. 2017, pp. 719–723.