International Journal for Modern Trends in Science and Technology
Volume 10, Issue 04, pages 321-325.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol10issue04.html
DOI: https://doi.org/10.46501/IJMTST1004048

Utilizing Machine Learning for Predicting Cardiovascular Stroke: An Advanced Prediction System

Dr.Manjula Devarakonda Venkata[1] ,Jami Rohit[2], Malakala Satya Priya[2], Yarrapragada Venkata Satya Aakash[2], Dachepalli Rajeshwar Pranay[2], Balireddy Siva Shankar Prasad[2]

[1]Associate Professor, Department of Computer Science Engineering, Pragati Engineering College, Surampalem , Andhra Pradesh, India.
[2]Department of Computer Science Engineering, Pragati Engineering College, Surampalem , Andhra Pradesh, India.

To Cite this Article
Dr.Manjula Devarakonda Venkata ,Jami Rohit, Malakala Satya Priya, Yarrapragada Venkata Satya Aakash, Dachepalli Rajeshwar Pranay, Balireddy Siva Shankar Prasad, Utilizing Machine Learning for Predicting Cardiovascular Stroke: An Advanced Prediction System, International Journal for Modern Trends in Science and Technology, 2024, 10(04), pages. 321-325. https://doi.org/10.46501/IJMTST1004048

Article Info
Received: 06 April 2024; Accepted: 18 April 2024; Published: 26 April 2024.

ABSTRACT

Cardiovascular illnesses have overtaken all other causes of death as the leading cause of death in industrialized, undeveloped, and developing countries in recent decades. Early identification of cardiac disease and therapeutic therapy can reduce the death rate. Based on the patient's numerous cardiac parameters, we suggested a model for forecasting heart illness and identifying imminent heart disease utilizing machine learning techniques such as logistic regression, SVM, Multinomial Nave Bayes, Random Forest, and Decision Tree. In most cases, input is provided in the form of numerical data representing various parameters, and output findings are created in real time to predict whether or not the patient has a disease. Before selecting which supervised machine learning method is ideal for the model, we'll experiment with several options.Existing systems use classical deep learning models, which are inefficient and imprecise. They are not as precise as the suggested model and take slightly longer to process.

Keywords: Machine learning, Black Box, Logistic regression, Heart ailment, SVM

1. INTRODUCTION

Machine learning is employed in a wide range of sectors worldwide. The healthcare industry is no different. Machine learning can be very beneficial in predicting whether or not locomotor problems, cardiovascular diseases, and other conditions will exist. If foreseen, such evidence can give clinicians with valuable information, allowing them to tailor their treatment plans and diagnoses. The heart is one of the body's primary organs. It forces blood to flow through the circulatory system's blood vessels. The circulatory system is necessary for supplying blood, oxygen, and other materials to the body's various organs. The heart is the most important part of the circulatory system. If the heart is not functioning properly, It can cause serious health complications, even death. Because the health care industry generates massive volumes of medical data, machine learning algorithms are critical for predicting

cardiac disease accurately. Recent research has focused on merging these approaches to develop hybrid machine learning algorithms. Data pre-processing is utilized in the study proposal to eliminate noisy data, fill in blanks as needed, fill in default values where suitable, and categorize attributes for prediction and decision-making at many levels. To evaluate the efficacy of the treatment strategy, techniques such as classification, accuracy, sensitivity, and specificity analysis are used. An accurate cardiovascular disease prediction model is presented by comparing the levels of accuracy while applying rules to the outcome variables of SVM Classifier, Decision Tree, Logical Regression, Random Forestand Multinomial Naïve Bayes.

2. LITERATURESURVEY

[1] Researchers use UCI repository data to test and train machine learning methods for predicting cardiac disease, including k-nearest neighbour, decision tree, linear regression, and support vector machine. The goal was to develop a method for predicting cardiovascular disease using classification techniques such as logistic regression, K-nearest neighbour, support vector machine, decision tree, and random forest, and to compare the effectiveness of the proposed scheme to the uniform standard.

In this work, [2] the authors present an overview of various categorization methods. Most classification approaches use benchmarks and threshold values to determine appropriate dataset properties.

[4] The Heart Disease Prediction System uses the MLP machine learning algorithm to assess a user's probability of developing CAD. The authors chose Multi-Layered Perceptron (MLP) as the recommended machine learning algorithm because to its efficiency and accuracy, resulting from recent technical advancements.

[5] This study uses machine learning to provide a uniquestrategy for predicting cardiovascular illnesses. To examinethe performance on the dataset, BettheAuthorsen's variousmachine learning methods are compared, and the proposed technique uses K Nearest Neighbour, which has an accuracy of 92.30 percent. According to machine learning, it has proven to be useful in generating inferences and forecasts from a large amount of data collected over time by the healthcare industry. [6] Artificial neural networks (ANN), decision trees (DT), random forests (RF), support vector machines (SVM), naïve Bayes (NB), and the k-nearest neighbour approach are among the machine learning techniques used to predict cardiac disease.Machine learning can predict a patient's prognosis by analysing medical records and identifying key details. To improve the performance of estimate procedures, it is critical to choose the best combination of relevant features. [7] The purpose of this study is to identify key characteristics and data mining approaches that can improve the accuracy of cardiovascular disease prediction.

The proposed diagnostic system is optimised using a grid search method. The Cleveland dataset, an online database of heart failure cases, is used in the investigations. Because the proposed method achieves 3.3 percent higher accuracy with only 7 features than the standard random forest model, it is both more effective and less difficult. The researchers who wrote this paper provide anovel method for employing machine learning Identifying key features improves the accuracy of cardiovascular disease prediction. The forecast model is introduced using several major features and numerous well-known categorization approaches.

.

3.SYSTEM ANALYSIS

A. EXISTING SYSTEM

The current approach for cardiovascular stroke prediction is based mostly on traditional deep learning models. These models are recognized for being inefficient and less exact than the proposed model. They often take longer to process and have lower accuracy. The existing system's input data is typically composed of numerical metrics relating to the patient's cardiac characteristics. The output is created in real time, and it predicts whether or not the patient has cardiac disease. The system makes these predictions using classic machine learning methods including logistic regression[14], SVM, Multinomial Naive Bayes, Random Forest, and Decision Tree. However, due to their inefficiencies and poorer accuracy, these models are less helpful in early identification and clinical management for heart problems.Critical in lowering the mortality rate caused by cardiovascular illnesses.

DISADVANTAGES OF THE EXISTING SYSTEM

1. **Inefficiency:** Classical deep learning models are computationally inefficient, which leads to longer processing times.
2. **Lower Precision:** The current system may have poorer precision and accuracy in predicting cardiac illnesses compared to more advanced machine learning techniques.
3. **Lack of Scalability:** Traditional models may struggle to scale well with huge datasets or integrate new data sources and features.
4. **Limited Feature Engineering:** The current system may have limited capabilities for feature engineering, which is crucial in capturing complicated relationships within the data.
5. **Reduced Adaptability:** These traditional models may not adapt effectively to shifting data patterns and new trends, rendering them unsuitable for real-time prediction and early detection of cardiac disease.

B. PROPOSED SYSTEM

In the proposed system, the machine learning model is trained in such a way that it can predict whether a person is at risk of developing cardiovascular disease. It forecasts and warns them of the impending danger. Prediction methods include Logistic regression, SVM, Naïve Bayes, and Machine Learning[17], which are commonly employed in classification problems. Based on the characteristics of the patient's heart, we should partition the data into different organized data sets.A logistic regression model is created to forecast the patient's disease based on the given data. The first step would be to collect data and preprocess it. The dataset is obtained from Kaggle, and its size should be huge in order to achieve high accuracy. Our dataset is very efficient in identifying and forecasting.

After data is acquired, it is preprocessed to eliminate all null values from the dataset. The Null values in the dataset will have a major impact on the training of the ML model. The dataset will contain a large amount of data, which must be normalized before training. Data transformation is the process of turning data into a format that is more favorable for data mining.

4. SYSTEM DESIGN

SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture.

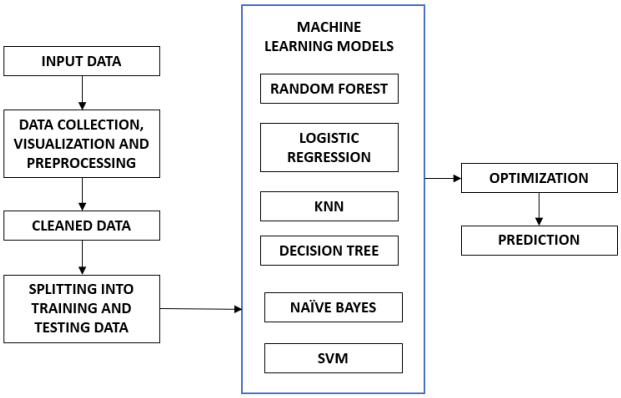

Fig. 1. Flowchart of proposed Model

5. SYSTEM IMPLEMENTATION

MODULES

Data Collection and Preprocessing: This module collects pertinent patient information such as medical history, vital signs, and test findings. It cleans, normalizes, and prepares the data for machine learning algorithms.

Feature Extraction and Selection: Feature extraction[15] reveals key traits from patient data that can be used to predict outcomes. Feature selection decides which features are most relevant to the model, lowering dimensionality and improving model performance.

Machine Learning Model Training: This module entails training and fine-tuning machine learning models such as logistic regression, SVM, Multinomial Naive Bayes, Random Forest, and Decision Tree on historical patient data to discover patterns and associations that can be utilized for prediction.

Real-time Prediction: Once the model has been trained, this module handles real-time predictions. It accepts new patient data and uses the trained model to forecast the possibility of cardiovascular illness, offering immediate feedback to healthcare practitioners.

Model Evaluation and Continuous Improvement: This module evaluates the performance of machine learning models using metrics such as accuracy, precision, recall, and the F1-score. It also enables model retraining with new data in order to react to shifting health patterns and enhance prediction accuracy.

6. RESULTS AND DISCUSSION

First, performance was measured using 10-fold cross validation in the training set. Second, results were obtained by simply applying the model; third, features were adjusted; and fourth, the framework was validated. This paper reviews current mining studies on cardiovascular disorders. Machine algorithms can be used to efficiently "mine" relevant data from the huge amounts of information generated by the medical business. Using a variety of mining algorithms produces considerably better results. A speedy and efficient deployment of cardiovascular disease monitoring is the result of a precise blend of mining algorithms and their application to the given dataset.

The essential dataset is divided into two halves, one for mining and the other for verification. Some studies examine multiple classification approaches on a dataset to determine whether a certain person is likely to develop cardiovascular disease. Others in the room have worked with data collection for a specific dataset of heart-related causes.In addition to examining these commonly used methodologies, several recent works have investigated "hybrid models". To get better outcomes, a hybrid model integrates many well-known classification and selection procedures into a single model. Hybrid models have been found to produce extremely accurate.

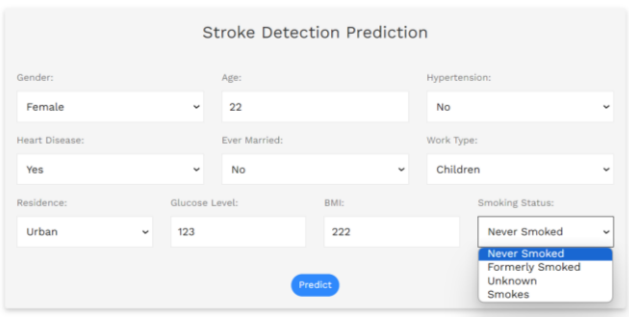

Fig 2. Giving report data as input to this application

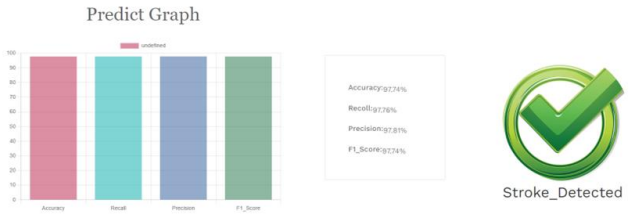

Fig 3.Strock Prediction

7. CONCLUSION AND FUTURE WORK

Early detection can help reduce illness progression and spread. Machine learning technology can help with early diagnosis and identifying crucial causal factors. The suggested technique produces a deep learning model capable of predicting cardiovascular diseases and heart attacks. The best solution for the job is the SVM algorithm. The model implies that new machine learning methods typically result in higher prediction accuracy. Early detection can help reduce illness progression and spread. Machine learning technology can help with early diagnosis and identifying crucial causal factors. The suggested technique produces a deep learning model capable of predicting cardiovascular diseases and heart attacks. The best solution for the job is the SVM algorithm.

The model implies that new machine learning methods typically result in higher prediction accuracy. Existing techniques are analyzed and compared to determine the most efficient and precise methods. Machine learning algorithms improve the accuracy of cardiovascular risk prediction, allowing people to be diagnosed early and get preventive medication. Some may claim that machine learning techniques are extremely promising for forecasting heart and circulation illnesses. All of the aforementioned strategies have performed admirably in all situations. With the multimodal technique, we were able to achieve a higher accuracy rate while also reducing processing time.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] A. Singh et al., "Heart Disease Prediction Using Machine Learning Algorithms", 2020 International Conference on Electrical and Electronics Engineering (ICE3), pp. 452-457, February 2020.

[2] K. Hashi and M.S.U. Zaman, "Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction",Journal of Applied Science & Process Engineering, vol. 7, no. 2, pp. 631-647, 2020.

[3] A. M.A and P. A. Thomas, "Comparative Review of Feature Selection and Classification modeling", 2019 International Conference on Advances in Computing Communication and Control (ICAC3), pp. 1-9, 2019

[4] A. Singh et al., "Heart Disease Prediction Using Machine Learning Algorithms", 2020 International Conference on Electrical and Electronics Engineering (ICE3), pp. 452-457, February 2020.

[5] K. Hashi and M.S.U. Zaman, "Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction",Journal of Applied Science & Process Engineering, vol. 7, no. 2, pp. 631-647, 2020.

[6] A. M.A and P. A. Thomas, "Comparative Review of Feature Selection and Classification modeling", 2019 International Conference on Advances in Computing Communication and Control (ICAC3), pp. 1-9, 2019

[7] A. Gavhane, G. Kokkula, I. Pandya and K. Devadkar, "Prediction of Heart Disease using Machine Learning", 2018 Second International Conference on Electronics Communication and Aerospace Technology (ICECA), pp. 1275-1278,2018.

[8] F. Mendonca, R. Manihar, A. Pal and S. U. Prabhu, "Intelligent Cardiovascular Disease RiskEstimation Prediction System", 2019 International Conference on Advances in Computing Communication and Control (ICAC3), pp. 1-6, 2019. 92 Authorized licensed use limited to: Universidad de Almeria. Downloaded on September 30,2023 at 14:01:46 UTC from IEEE Xplore. Restrictions apply. 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)

[9] Rahul Katarya and PolipireddySrinivas, "Predicting heart disease at early stages using machine learning: A survey", 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020.

[10] DavideChicco and Giuseppe Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone", BMC medical informatics and decision making, vol. 20.1, pp. 1- 16, 2020.

[11] Mohammad Shafenoor Amin, Yin KiaChiamandKasturiDewiVarathan, "Identification of significant features and data mining techniques inpredicting heart disease", Telematics andInformatics, vol. 36, pp. 82-93, 2019

[12] AshirJaveed et al., "An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection", IEEE Access, vol. 7, pp.180235-180243, 2019

[13] S. Mohan, C. Thirumalai and G. Srivastava,"Effective heart disease prediction using hybrid machine learning techniques", IEEE access, vol.7, pp. 81542-81554, 2019

[14] Manjula Devarakonda Venkata, Sumalatha Lingamgunta," Triple-Modality Breast Cancer Diagnosis and Analysis in middle aged women by Logistic Regression", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-4, February 2019

[15] Manjula Devarakonda Venkata, Sumalatha Lingamgunta, "Statistical and Texture Features based Breast Cancer Classification from Multi image Modalities using Support Vector Machines with Gaussian Radial Basis Function", International Journal of Control and Automation Vol. 13, No.1, (2020), pp. 384 - 396

[16] Shaik Mohammad Rafee, Manjula Devarakonda Venkata," An Ensemble Neuro Fuzzy Algorithm For Breast Cancer Detection And Classification", Ictact Journal On Soft Computing, January 2024, Volume: 14, Issue: 03, DOI: 10.21917/ijsc.2024.0460

[17] Manjula Devarakonda Venkata, Sumalatha Lingamgunta Health Care Automation in Compliance to Industry 4.0 Standards :A Case Study of Liver disease Prediction", Journal of Scientific & Industrial Research Vol. 82, February 2023, pp. 263-268 DOI: 10.56042/jsir.v82i2.70215