IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

# A Machine Learning Approach For Cardiovascular Stroke Prediction System

**1 Chikkireddy Sri Ravali 2 M Subrahmanyeswara Rao 3 V Anil Santosh**

1 M.Tech Scholar, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India, sriravalidegala@gmail.com.

2 Associate professor, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India.

3 Associate Professor and HoD, Department of CSE, International School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India.

**ABSTRACT:** Over the past few decades, cardiovascular diseases have surpassed all other causes of death as the main killers in industrialized, underdeveloped, and developing nations. Early detection of heart conditions and clinical care can lower the death rate. Based on the patient's various cardiac features, we proposed a model for forecasting heart disease and identifying impending heart disease using machine learning techniques such logistic regression, SVM, Multinomial Nave Bayes, Random Forest, and Decision Tree. In most cases, input is received through numerical data of various parameters, and output findings are generated in real-time, predicting whether or not the patient has a disease. We'll use a variety of supervised machine learning methods before deciding which one is best for the model. Existing systems rely on classical deep learning models, which are inefficient and imprecise. They aren't as accurate as the proposed model and take a little longer to process.

**1. INTRODUCTION:** Around the world, machine learning is used in many different industries. The healthcare industry is no exception. Machine learning can be quite useful in determining whether or not ailments like locomotor disorders, cardiovascular diseases, and others will exist. If foreseen, such evidence can provide clinicians with insightful knowledge, enabling them to customize their treatment plans and diagnoses. One of the body's major organs is the heart. It forces blood through the blood vessels of the circulatory system. The body's numerous organs are supplied with blood, oxygen, and other materials by the circulatory system, which is essential. The heart is the most important component of the circulatory system. If the heart does not operate properly, it can result in major health problems, including death. The health care sector has vast amounts of medical data; consequently, machine learning algorithms are essential for accurate heart disease prediction. Recent studies have focused on combining these strategies to create hybrid machine learning algorithms. Data pre-processing is used in the research proposal to remove noisy data, fill in blanks when necessary, fill in default values when appropriate, and categorize attributes for prediction and decision-making at multiple levels. To assess the efficacy of the treatment approach, techniques including classification, accuracy, sensitivity, and specificity analysis are performed. An accurate cardiovascular disease prediction model is demonstrated by comparing the levels of accuracy of applying rules to the outcome variables of SVM Classifier, Decision Tree, Logical Regression, Random Forest, and Multinomial Naïve Bayes.

## 2 SYSTEM ANALYSIS:

### 2.1 EXISTING SYSTEM:

The existing system for cardiovascular stroke prediction relies primarily on classical deep learning models. These models are known to be inefficient and less precise compared to the proposed model. They often require more time for processing and have lower accuracy. The input data in the existing system typically consists of numerical parameters related to the patient's cardiac features. The output is generated in real-time and provides predictions regarding whether or not the patient has a heart disease. The system uses traditional machine learning algorithms such as logistic regression, SVM, Multinomial Naive Bayes, Random Forest, and Decision Tree to make these predictions. However, the inefficiencies and lower accuracy of these models make them less effective in early detection and clinical care for heart conditions, which are critical in reducing the death rate associated with cardiovascular diseases.

LIMITATIONS

Inefficiency: Classical deep learning models tend to be computationally inefficient, resulting in longer processing times.

Lower Precision: The existing system may have lower precision and accuracy in predicting heart diseases compared to more advanced machine learning techniques.

Lack of Scalability: Traditional models may struggle to scale effectively with large datasets or incorporate new data sources and features.

Limited Feature Engineering: The existing system may have limited capabilities for feature engineering, which can be critical in capturing complex relationships within the data.

Reduced Adaptability: These classical models may not adapt well to changing data patterns and emerging trends, making them less suitable for real-time prediction and early detection of heart conditions.

### 2.2 PROPOSED SYSTEM:

In the proposed system, the machine learning model is trained in such a way that the system could predict if the person has the chance of getting cardio vascular disease. It predicts and alerts them about the impending danger. The method used for prediction uses Logistic regression, SVM, Naïve Bayes, Machine Learning algorithms which are widely used in classification-based problem. Based on the characteristics of the patient's heart, the data we have should be divided into several structured data sets.

Using a logistic regression approach, a model is developed that predicts the patient's illness based on the available data. The initial step would be to collect data and then preprocess them. The dataset is collected form the Kaggle and the size of the dataset should be large to have high accuracy and our dataset is very efficient in detecting and predicting.

After data is collected, they are preprocessed to remove all the null values in the dataset. The Null values in the dataset will have significant effect on the training of the ML model. The dataset will have a lot of data which needs to be normalized before they are trained. Data transformation is the process of converting data into a format that is more conducive to data mining.

## 3. SYSTEM IMPLEMENTATION:

**MODULES**

**Data Collection and Pre-processing**

**Feature Extraction and Selection**

**Machine Learning Model Training**

**Real-time Prediction**

**Model Evaluation and Continuous Improvement**

**Data Collection and Preprocessing:**

This module is responsible for collecting relevant patient data, such as medical history, vital signs, and lab results. It pre-processes the data to clean, normalize, and prepare it for machine learning algorithms.

**Feature Engineering and Selection**

Feature extraction identifies important characteristics from the patient data that can be used for prediction. Feature selection determines which features are most relevant for the model, reducing dimensionality and enhancing model performance.

**Model Development**

This module involves 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 learn patterns and relationships that can be used for prediction.

**Real-Time Monitoring and Alerting**

Once the model is trained, this module handles real-time predictions. It takes in new patient data and uses the trained model to predict the likelihood of cardiovascular disease, providing instant feedback to healthcare professionals.

**Explainability and Adaptation**

This module assesses the performance of the machine learning models using metrics like accuracy, precision, recall, and F1-score. It also facilitates model retraining with new data to adapt to changing health trends and improve predictive accuracy.

## SYSTEMREQUIREMENTS

### 3.1 HARDWARE REQUIREMENTS

| | | |
|---|---|---|
| System | : | Intel i3 2.2 Ghz |
| RAM | : | 4GB |
| Hard Disk | : | 256GB |

### 3.2 SOFTWARE REQUIREMENTS

| | | |
|---|---|---|
| OS | : | Windows 10 or 11 |
| Programming Language | : | Python |
| Tool | : | VS code |

### SOFTWARE ENVIRONMENT

### PYTHON

Python is a High level, structured, open-source programming language that can be used for a wide variety of programming tasks.

Python within itself is an interpreted programming language that is automatically compiled into bytecode before execution.

It is also a dynamically typed language that includes (but does not require one to use) object-oriented features.

NASA has used Python for its software systems and has adopted it as the standard scripting language for its Integrated Planning System.

Python is also extensively used by Google to implement many components of its Web Crawler and Search Engine & Yahoo! for managing its discussion groups

## 4. SYSTEM TEST:

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

**4.1 TYPES OF TESTS:**

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screensor fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manual

**CONCLUSION:** The prevention and disease progression can be aided by early identification. Early diagnosis and the finding of important causal factors can be aided by machine learning technologies. The proposed technique generates a deep learning model that can predict cardiovascular illnesses and heart attacks. The optimal solution for the job is the SVM algorithm. The model suggests that new machine learning algorithms usually lead to better prediction accuracy. The prevention and disease progression can be aided by early identification. Early diagnosis and the finding of important causal factors can be aided by machine learning technologies. The proposed technique generates a deep learning model that can predict cardiovascular illnesses and heart attacks. The optimal solution for the job is the SVM algorithm. The model suggests that new machine learning algorithms usually lead to better prediction accuracy. For the purpose of finding efficient and precise methods, existing techniques are evaluated and compared. Machine learning algorithms greatly enhance the accuracy of cardiovascular risk prediction, allowing individuals to be detected early in the disease process and get preventive treatment. One may argue that machine learning techniques are very promising for predicting heart- or circulatory related diseases. All of the aforementioned techniques have performed superbly in every situation. With the multimodal strategy, we were able to obtain a greater accuracy rate while also reducing processing time.

**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 Polipireddy Srinivas, "Predicting heart disease at early stages using machine learning: A survey", 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020.

[10] Davide Chicco 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 Kia Chiamand Kasturi Dewi Varathan, "Identification of significant features and data mining techniques inpredicting heart disease", Telematics andInformatics, vol. 36, pp. 82-93, 2019.

[12] Ashir Javeed 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.



**Biography of authors:**

**Chikkireddy Sri Ravali** was a M.Tech Scholar in Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. Her intersted research area  is machine learning and artificial intelligence (AI) typically focuses on advanced computational techniques.

**M Subrahmanyeswara Rao** was a  Associate Professor (PhD) in Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. His current research work is  is machine learning and artificial intelligence (AI) typically focuses on advanced computational techniques that enable machines to learn from data, identify patterns, and make decisions without being explicitly programmed.

**V Anil Santosh** was a Associate Professor(PhD) and Head of the Department of Department of CSE inInternational School Of Technology And Sciences For Women(A) NH-16 East Gonagudem Rajanagaram, AP, India. His current research work is a variety of AI subfields, including deep learning, neural networks, natural language processing, and reinforcement learning. Their work may involve developing new algorithms, applying AI to solve real-world problems (like forecasting, automation, or image recognition), and exploring ethical concerns related to AI deployment. Many such authors combine academic research with industry applications, publishing papers, books, or articles aimed at both technical and non-technical audiences