iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.53233

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Comparative Study to Identify the Heart Disease Using Machine Learning Algorithms

Aditya Khandare[1], Pavan Sirsath[2], Gaurav Borse[3], Rutuja Nale[4], Tejasvi Navale[5], Satish Gujar[6]
[1, 2, 3, 4, 5]*B.E. Student, Tssm's Bhivarabai Sawant College Of Engineering and Research, Pune (Maharashtra), India*
[6]*Professor, Tssm's Bhivarabai Sawant College Of Engineering and Research, Pune (Maharashtra), India*

Abstract: ***Heart disease is a major health concern globally, and it is a leading cause of death, especially in developing countries in Africa and Asia. Early detection of heart disease can play a crucial role in preventing its occurrence and reducing its impact. This is where the use of deep neural networks (CNNs) and the Django framework can be highly beneficial. The CNN algorithm is a machine learning technique that uses multiple layers of artificial neural networks to analyze complex data and identify patterns. In the case of heart disease prediction, the CNN algorithm can be trained on a dataset such as the UCI dataset, which contains a large amount of data related to heart disease risk factors and patient characteristics. The dataset can be split into training and testing data to enable the algorithm to learn and validate its accuracy.***
Keywords: Django framework, globally, dataset, artificial neural networks, machine learning, identify patterns.

I. INTRODUCTION

Deep learning is a powerful subset of machine learning that enables artificial neural networks to learn from vast amounts of data. Unlike traditional machine learning methods that require human engineers to design and engineer features, deep learning algorithms can automatically learn and extract relevant features from raw data. This makes deep learning particularly well-suited for analyzing unstructured data, such as images, audio, and natural language.

One of the most well-known applications of deep learning is in computer vision, where convolutional neural networks (CNNs) have revolutionized the field. CNNs can detect and classify objects in images and video with remarkable accuracy, and are used in a wide range of applications, from self-driving cars to medical image analysis.

Another area where deep learning is having a significant impact is in natural language processing (NLP). Recurrent neural networks (RNNs) and their variants, such as long short-term memory (LSTM) networks, are used to analyze and generate natural language text. This has enabled major advances in areas such as language translation, chatbots, and voice assistants.

Generative adversarial networks (GANs) are another type of neural network used in deep learning, which are used to generate synthetic data that is similar to real-world data. This has applications in fields such as art, music, and design, where GANs can be used to generate new, creative content.

Overall, deep learning has the potential to revolutionize many industries and is already having a significant impact on the field of artificial intelligence. As the amount of data we generate continues to grow exponentially, deep learning will become even more important for analyzing and extracting insights from this data. The possibilities are virtually endless, and we are only beginning to scratch the surface of what is possible with this exciting technology.

II. PROPOSED SYSTEM

The proposed system aims to predict the risk of cardiovascular disease using a deep neural network (CNN) approach. Cardiovascular disease is a condition that affects the heart and blood vessels, and it is one of the leading causes of death worldwide. The major cause of heart attack in the United States is coronary artery disease, which is caused by the narrowing of coronary arteries that supply blood to the heart. The incidence of cardiovascular disease is higher in men than women, and it is estimated that 24 percent of deaths in India are due to this condition.

In practical applications, neural networks have shown high accuracy in predicting cardiovascular disease. The dataset is divided into two parts: the training dataset and the testing dataset. The training dataset is used to train the neural network, while the testing dataset is used to evaluate the performance of the trained model. Neural networks are a set of algorithms used to recognize patterns in data, and they consist of layers of neurons or nodes. Each node in the network is connected to the nodes in the previous and next layers by weighted connections.

	Predicted heart disease patient	Predicted healthy patient
Actual heart disease patient	TP	FN
Actual healthy patient	FP	TN

The training features are fed to the input layer, and the actual processing happens in the hidden layers using activation functions. The output layer of the network is connected to the hidden layer, and the aim of the predictive model is to generate a hypothesis through deep learning models. The hypothesis is the relationship between data, which can be tested by collecting data and making observations. The model generates the hypothesis by minimizing the error in the training instances.

$$f(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

The performance of the network is dependent on the number of guidelines used, which decide the behavior of the network. A model with fewer parameters leads to low capacity, which results in
underfitting. A model with more parameters than required leads to high capacity, which results in overfitting. The model should be designed in such a way that it generates a hypothesis with optimum capacity.
The hypothesis is formulated using forward propagation. The input is given to the neurons, which perform some operation to generate the output. This process is called the activation function. The activation function defines the output of a node. Deep neural networks contain more than one hidden layer, and the activation of neurons present at the output layer is achieved using the sigmoid activation function. Feature selection is used to filter out redundant features from the dataset. Feature extraction is different from feature selection, as it involves getting useful elements from the existing data. Relevant elements are fed to the neural network by eliminating the irrelevant features using feature selection. Irrelevant elements are removed using two statistical models. The dependence between the feature and class is measured using two tests. Features are ranked in the first step, and from the ranked features, optimal features are searched in the second step.
Heart disease binary classification contains two classes: the positive class and the negative class. It also contains t instances, and can represents the expected values. The expected value of two independent value based on null hypothesis can be calculated.

$$E_a=(a+b)\ \frac{a+b}{t} \qquad (2)$$

$$x^2 = \frac{1}{d}\sum_{k=1}^{n}\frac{(O_K-E_K)^2}{E_K} \qquad (3)$$

The threshold for the number of features must be decided after feature ranking, which is denoted by n. The subset of features with n=1 is taken, and the optimum number of features is found through exhaustive search. The subset of features is applied to CNN, and grid search is used to evaluate the CNN performance. After saving the result of the first subset, another subset with n=2 is taken, and the optimum feature is found. Then it is applied to CNN, and the result is saved. These procedures are continued all the unit.

III. SYSTEM ARCHITECTURE

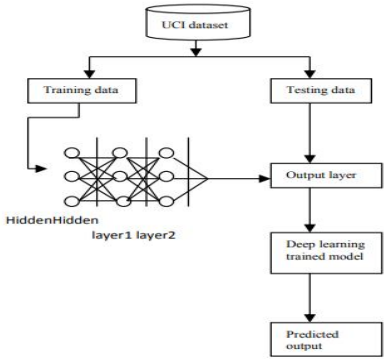


Fig 1: System Architecture

A. *Basics of CNN*

One of the key features of a great Content Management System (CMS) is its ability to make content management easy and intuitive for users, regardless of their technical background. Dot.Net.Nuke (CNN) is an excellent CMS that achieves this goal by providing a user-friendly interface and a rich set of tools for content creation and editing.

The rich text editor in CNN allows users to create and format content in a visually appealing way. The editor includes a variety of formatting options, such as bold, italic, underline, bullet points, and numbering, as well as the ability to add images, videos, and provide accurate and cost-effective solutions for disease prediction.

other multimedia content. Additionally, CNN offers many options for creating custom content types, such as blog posts, news articles, product descriptions, and more.

For larger teams with multiple content contributors, CNN offers advanced workflow management features. Content workflows can be set up to ensure that content is properly reviewed and approved before it is published. Users can be assigned different roles and permissions based on their level of involvement in the content creation and approval process.

For example, a manager might have the ability to approve or reject content, while an editor might only be able to create and edit content. Overall, CNN's content management features make it an ideal CMS for businesses of all sizes that need to create and manage content efficiently and effectively.

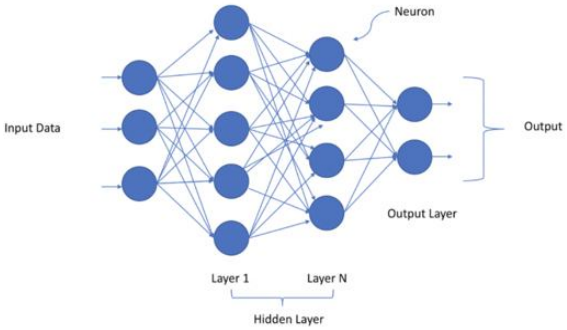


Fig 2: CNN Algorithm

IV. LITERATURE SURVEY

In Dengqing Zhang et.al's paper, a novel heart disease prediction model is proposed. The model combines an embedded feature selection method and deep neural networks to improve the accuracy of heart disease prediction. The embedded feature selection method is based on the Linear SVC algorithm, which uses the L1 norm as a penalty item to select a subset of features significantly associated with heart disease. The selected features are then fed into the deep neural network built by the authors. The weights of the network are initialized with the He initializer, which helps prevent gradient vanishing or explosion and improves the predictor's performance. To test the model's effectiveness, the authors used the heart disease dataset obtained from Kaggle. The dataset includes various attributes related to heart disease, such as age, sex, blood pressure, cholesterol levels, and more. The authors trained their model on this dataset and achieved promising results, demonstrating the effectiveness of their proposed algorithm. Overall, Dengqing Zhang et.al's paper presents a promising approach to predicting heart disease using deep neural networks and embedded feature selection methods. Their work highlights the importance of accurately selecting relevant features and initializing the network's weights to improve the predictor's performance.

P Kalpana1 et.al's paper emphasizes the importance of early detection and prevention of coronary heart disease. The authors argue that it is crucial to pay attention to the signs of heart disease in the early stages to prevent the high mortality rate associated with it. They suggest that individuals should undergo more frequent electrocardiogram (ECG) tests to detect any abnormalities in the heart's functioning. The authors propose the development of a nursing assistant system that can predict the risk of coronary heart disease based on key indicators such as age, course, and heartbeat. They suggest that neural codes for learning are the most reliable and robust and, therefore, should be used to develop the system. Overall, P Kalpana1 et.al's paper highlights the need for early detection and prevention of coronary heart disease to reduce the high mortality rate associated with it. Their proposed nursing assistant system could potentially help identify individuals at risk of developing heart disease, allowing for early intervention and prevention. The authors' focus on using neural codes for learning reflects the growing importance of artificial intelligence and machine learning in healthcare

In the paper by Awais Mehmood et.al [3], the authors introduce a system called Cardio Help, which is designed to predict the probability of the presence of cardiovascular disorder in a patient. The system uses a deep learning approach known as convolutional neural networks (CNNs) to analyze time-series data and identify potential cases of heart failure at an early stage. The authors tested the system on a dataset of patients with coronary disease and compared the results with those obtained using other approaches.

The initial results showed that Cardio Help outperformed other existing methods in terms of performance measures. The system has the potential to help medical practitioners in diagnosing heart disease more accurately and at an early stage, which could lead to better treatment outcomes and improved patient care.

In their research [4], Ufaq Jeelani Khan et al. proposed a hybrid scheme for heart disease prediction, which involved using random forest (RF) and decision tree (DT) algorithms. The first step was to extract the features using the RF algorithm, which was applied to the dataset containing various heart disease attributes.

The extracted features were then used to implement the DT algorithm for classification purposes. The proposed model was evaluated through performance analysis and it was found that the recommended model achieved an accuracy of approximately 94.44 percent. This indicates that the hybrid scheme can be effective in predicting heart disease, as it provides a high level of accuracy, precision, and recall. These findings are valuable for medical practitioners and researchers who are interested in developing predictive models for heart disease.

In this paper [5], the authors focus on the importance of using data mining and machine learning techniques to extract valuable information from datasets related to heart diseases. They highlight the fact that there are several available techniques that can be used to extract relevant information from the dataset by analyzing various attributes related to heart disease. The primary goal of this paper is to provide a summary of recent research in the field of heart disease prediction and to build scientific conclusions by using data mining and classification techniques through machine learning. By applying these techniques to the dataset, the authors aim to extract valuable insights that can aid in predicting the likelihood of heart disease in patients. data mining and machine learning techniques to extract information from datasets, particularly when it comes to predicting heart disease. The authors argue that these techniques can help medical professionals make more informed decisions about diagnosis and treatment, ultimately leading to better outcomes for patients

In this paper [6], the authors aimed to improve the accuracy levels of different AI techniques for heart disease prediction. They utilized a dataset collected from Kaggle that includes various attributes related to coronary disease such as age, gender, blood pressure, cholesterol, etc. The authors analyzed the accuracy levels of different AI techniques such as Support Vector Machines (SVM), K-Nearest Neighbor (KNN), and Decision Trees (DT). However, the results and accuracy of these techniques were not satisfactory when applied to a large dataset. Therefore, the authors proposed using Artificial Neural Network (ANN) and Tensor Flow Keras to improve the accuracy levels. ANN is a machine learning technique inspired by the structure and function of the human brain. It consists of multiple layers of interconnected nodes or neurons that are capable of learning complex patterns and relationships from the input data.

The authors implemented the proposed model using Tensor Flow Keras and compared its performance with other techniques. The results showed that the proposed model using ANN and Tensor Flow Keras achieved higher accuracy levels compared to other techniques, thus demonstrating the effectiveness of the proposed approach for heart disease prediction. In this paper[7] , the authors propose a model for predicting the likelihood of an individual having a coronary disease using K-Nearest Neighbor (KNN) and Logistic Regression classifiers. They compared the accuracy of their proposed model with the commonly used classifiers like Naive Bayes and found their model to be more accurate. The proposed model reduces the cost of clinical care and provides valuable insights to help predict patients with coronary disease. The model was implemented using Python and the results showed promising accuracy, making it a useful tool in the diagnosis and prevention of coronary disease. Overall, the study highlights the potential of machine learning in healthcare and its ability to provide accurate and cost-effective solutions for disease prediction.

V. SOME OF THE ADVANAGES FROM THE ABOVE RESULTS:

1) Increased accuracy for effective heart disease diagnosis:
2) With the use of deep neural algorithms and feature selection techniques, heart disease detection systems can provide a more accurate diagnosis than traditional methods.
3) By analyzing a vast amount of data, the system can identify potential risk factors and predict the likelihood of heart disease with greater accuracy Time-saving for doctors:
4) The heart disease detection system can help doctors save time by automating many tasks, such as data analysis, pattern recognition, and decision making.
5) This can help doctors to focus more on patient care and treatment, leading to better outcomes for patients Cost-effective for patients.
6) By providing accurate and timely diagnosis, the heart disease detection system can help improve patient outcomes.

VI. CONCLUSION

1) In addition to using different machine learning algorithms, future work can explore the use of deep learning techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for heart disease detection.
2) Further research can be done on incorporating additional medical parameters and demographic information for more accurate predictions.
3) The system can be further developed to provide personalized treatment plans for patients based on their individual health data and medical history.
4) Future work can also focus on developing a mobile application for heart disease detection, which can be more accessible and convenient for patients.
5) Research can be done to determine the efficacy and feasibility of implementing this system in real-world clinical settings.
6) Finally, the system can be extended to include other types of cardiovascular diseases such as stroke, hypertension, and peripheral artery disease.

REFERENCES

[1] Dengqing Zhang,1,2 Yunyi Chen,3 Yuxuan Chen,3 Shengyi Ye,1,2 WenyuCai,1,2 Junxue Jiang,1,2 Yechuan Xu,1,2 Gong feng Zheng,1,2 and MingChen1,4 "Heart Disease Prediction Based on the Embedded Feature Selection Method and Deep Neural Network" Received21 Aug 2021 Revised07 Sep2021 Accepted16 Sep 2021 Published29 Sep 2021.
[2] P Kalpana1 , S Shiyam Vignesh1 , L M P Surya1 , V Vishnu Prasad "Prediction of Heart Disease Using Machine Learning "Journal of Physics: Conference Series 1916 (2021) 012022.
[3] Awais Mehmood1 • Munwar Iqbal1 • Zahid Mehmood2 • Aun Irtaza1 •Marriam Nawaz1 • Tahira Nazir1 • Momina Masood1 "Prediction of Heart Disease Using Deep Convolutional Neural Networks "Received: 6 May 2020/ Accepted: 2 November 2020 © King Fahd University of Petroleum and Minerals 2021.
[4] S. Verma and A. Gupta, "Effective Prediction of Heart Disease Using Data Mining and Machine Learning: A Review," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021.
[5] U. J. khan, A. oberoi and J. Gill, "Hybrid Classification for Heart Disease Prediction using Artificial Intelligence," 2021 5th International Conference on XX VIII Computing Methodologies and Communication (ICCMC), 2021.

Software Requirements

[1] Operating System : Windows, Linux or MacOS
[2] Programming Language : Advance Java
[3] Tool : Android Studio
[4] Database : MYSQL Lite , CSV File

BIOGRAPHIES

Aditya Khandare
B.E. Student , Dept. of Computer Engineering

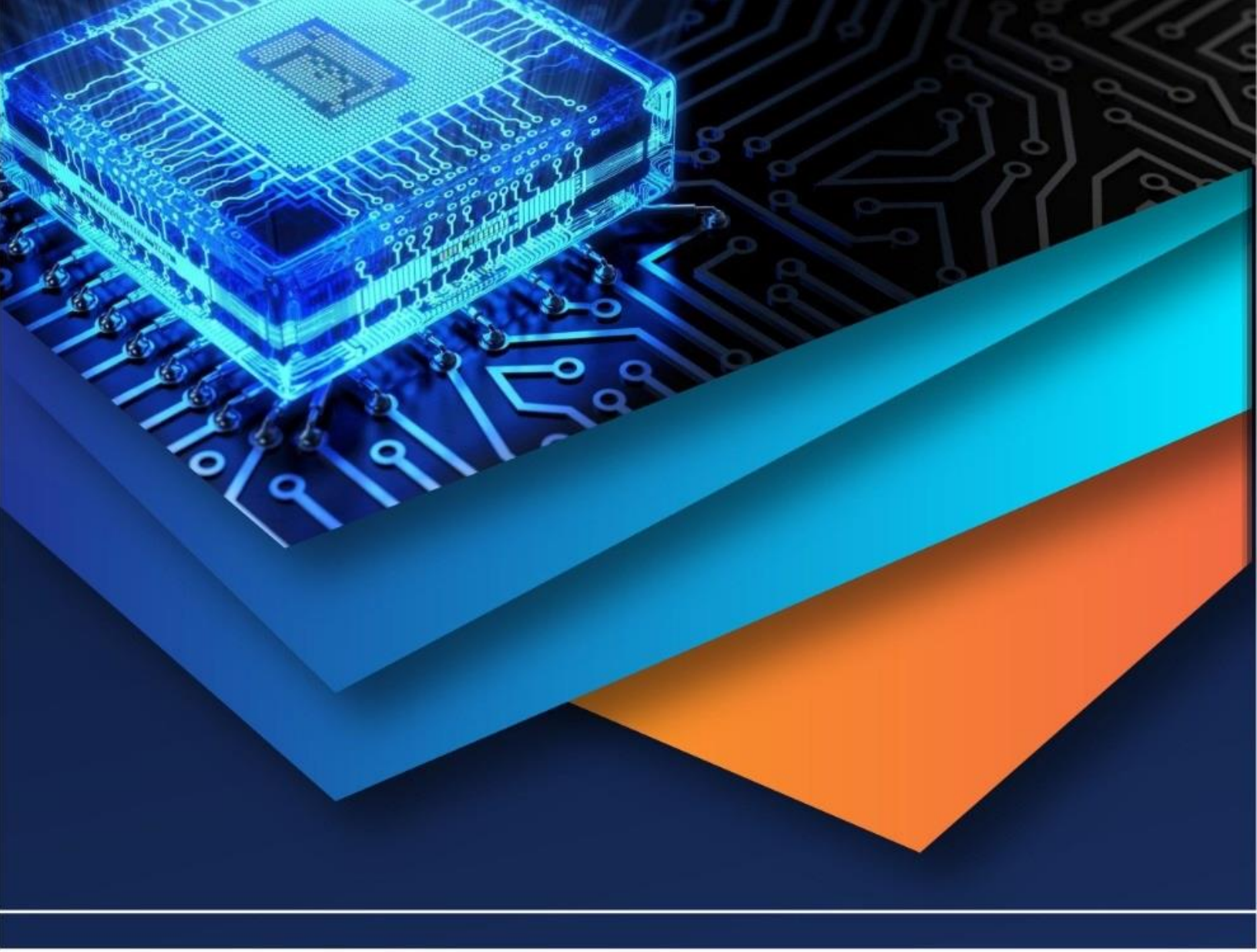
Pavan Sirsath
B.E. Student , Dept. of Computer Engineering

Gaurav Borse
B.E. Student , Dept. of Computer Engineering

Rutuja Nale
B.E. Student , Dept. of Computer Engineering

Tejasvi Navale
B.E. Student, Dept. of Computer Engineering

Prof. D.r. Satish Gujar
Head of Dept, Dept. of Computer Engineering



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY