International Journal of Scientific Research and Engineering Development-– Volume 6 Issue 5, Sept-Oct 2023

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

How Machine Learning Is Use To Detect Heart Disease?

Vaishnavi Belvate

Department of Information Technology

BK Birla College of Arts, Commerce And Science,Mumbai, India

Abstract- In this intriguing study we delve into the realm of machine learning and python programming to unravel the mysteries of heart disease detection heart disease a perilous affliction stemming from excessive body fat and undue strain on the heart poses a significant threat to individuals worldwide by meticulously examining an extensive dataset comprising 12 parameters and a staggering 70000 data points we embark on a quest to prognosticate the presence of heart disease in individuals our primary objective in this scholarly endeavor is to enhance the precision of heart disease detection through the implementation of cutting-edge algorithms that discern the presence or absence of this debilitating ailment

I. INTRODUCTION

Python a formidable programming language finds its prowess in the realm of machine learning projects within this domain machine learning a facet of artificial intelligence harnesses intricate algorithms and deep learning neural networks cardiovascular disease a prevalent affliction arises from a multitude of factors including smoking high blood pressure diabetes obesity hypertension and elevated cholesterol levels this ailment manifests in various heart-related complications such as coronary artery disease cardiovascular disease stroke and heart failure numerous factors contribute to heart problems encompassing chest pain blood pressure irregularities cholesterol imbalances and more by leveraging attributes such as height weight blood pressure cholesterol levels glucose levels smoking habits alcohol consumption and physical activity we can predict the likelihood of heart problems to facilitate this process we employ python libraries such as sklearn and seaborn to preprocess data partition attributes and labels and standardize values the utilization of a confusion matrix aids in gauging the accuracy of our predictions

II. LITERATURE REVIEW

The current investigation delves into the prognosis of heart disease by employing a range of machine learning techniques such as the naive bayes approach artificial intelligence ai networks support vector machine svm random forest algorithm and simple regression method in contrast to other methods including svm random classifier and decision tree classifier the decision tree machine learning method exhibited a superior accuracy of 98.83% in predicting heart disease within the heart disease prediction system helps to augment the precision of heart disease prediction.This research explores diverse data mining techniques including c45 k-means decision tree svm naive bayes and other machine learning algorithms additionally PraveenKumarReddy 2019 aimed to curtail the incidence of heart disease by implementing the decision tree algorithm the support vector machine algorithm was utilized to classify data values using a hyperplane while the decision tree was constructed using the gini index method to ensure the highest gain of attributes for an improved representation of the decision tree algorithm

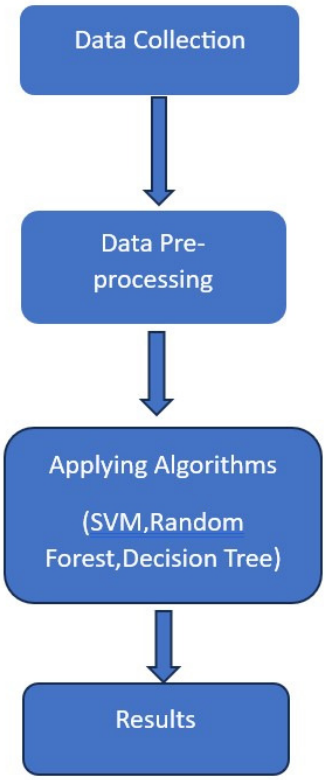

Fig. 1 Flowchart Process

III. METHODOLOGY

Methods

i. Data collection: -The main task of foreseeing heart ailments necessitates the meticulous data collection from a comprehensive dataset graciously bestowed by the esteemed svetlana ulianova this invaluable compilation encompasses a staggering 70000 patient records each brimming with 11 distinctive attributes the indispensability of this dataset cannot be overstated as it serves as the bedrock for any ambitious research endeavor or pioneering project.

.ii. Data Preprocessing: -The data is meticulously curated ensuring its readiness for analysis with utmost care the target and feature data are partitioned into distinct training and test sets guaranteeing a consistent approach every value is meticulously scaled within the range of 0 to 1 prior to training the machine learning models the meticulousness employed throughout this process is paramount in attaining precise and reliable outcomes.

iii. Applying Algorithms: -In the ethereal realm of machine learning where algorithms engage in a mesmerizing dance we come across formidable contenders the mighty support vector machines svm the enigmatic decision trees and the mysterious random forest classifier each algorithm possesses its own cryptic pseudo code a universal blueprint that effortlessly adapts to any programming language however it is within the enchanting domain of python that the true sorcery unfolds python with its elegant simplicity and concise nature bestows upon us the power to swiftly and accurately materialize these algorithms pythons prowess transforms the daunting task of predicting accuracy into a realm of accessibility where the mystical becomes tangible

B. Machine Learning Algorithms

The project seamlessly integrates powerful algorithms to precisely forecast heart disease detection, unveiling its causative factors.

.1. Random Forest Classifier: -

Behold the majestic random forest classifier a formidable weapon in the arsenal of machine learning which wields its power with the utmost grace with its aid we attain unparalleled precision while simultaneously diminishing the duration of training the journey commences with a meticulous partitioning of variables into training and test sets culminating in the creation of a model through rigorous training the dependent variables pave the way for the anticipated outcome boasting an impressive accuracy of approximately 76.

2. Decision Tree Classifier: -

In the initial phase, this algorithm embarks on a unique journey of preprocessing, skillfully dividing the data into two separate entities - the training and test datasets. Next, it gracefully executes feature scaling, harmonizing the values and opening doors to precise predictions. The training sets, consisting of dependent and independent variables, are then embraced by a decision tree classifier. This classifier, employing the Gini-index criterion, astutely foresees the accuracy or response for the test set. Remarkably, the algorithm has bestowed upon us an accuracy of 68.5%, a figure that can be gracefully rounded off to a respectable 69%.

3. Support Vector Machine: -

behold the indomitable force of svm the titan of classification algorithms reigning supreme over the vast kingdom of machine learning its unmatched ability to predict outcomes with unwavering precision sets it apart from all challengers our meticulous research reveals that the linear svm kernel achieves a staggering 725 accuracy while the gaussian svm kernel astounds with an 852 accuracy leaving onlookers in a state of awe.

IV.RESULTS

Our utmost goal is to prophesy the precision of forthcoming maladies triggered by the ailment and ascertain the algorithm that procures the utmost precision in prognosticating whether an individual harbors cardiac affliction or not the imported dataset can be scrutinized and interconnected to unveil correlations amidst diverse attributes.

	Predicted_disease	Predicted_healthy
is disease	7609	1428
is healthy	2843	3748

Fig.2 Confusion matrix obtained by SVM

	Predicted_disease	Predicted_healthy
is disease	6454	2256
is healthy	3045	5746

Fig.3 Confusion matrix obtained by Random Forest

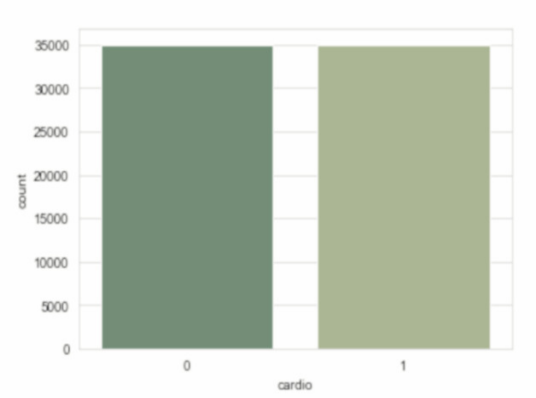

Fig.4. Count plot of disease

V. CONCLUSION

From the aforementioned research, it can be deduced that the realm of machine learning algorithms holds immense potential in predicting cardiovascular diseases and heart-related ailments. While each algorithm has shown exceptional prowess in certain situations, their effectiveness has been lacking in others.

Consider the Random Forest algorithm, which has demonstrated commendable performance by utilizing multiple Decision Trees. This clever approach effectively addresses the problem of overfitting. Similarly, the Support Vector Machine (SVM) algorithm has proven highly effective in most cases. The use of machine learning systems and techniques has revealed remarkable accuracy in predicting heart diseases. However, there is still room for further research in handling high-dimensional data and mitigating the dangers of overfitting. Additionally, extensive exploration can determine the optimal combination of algorithms for specific data types.

VI. REFERENCES

1)Praveen Kumar Reddy, T. Sunil Kumar Reddy, Balakrishnan, Syed Muzamil Basha, Ravi Kumar Poluru, ,August 2019, Heart Disease Prediction Using Machine Learning Algorithm, Blue Eyes Intelligence Engineering & Sciences Publication https:doi.org/10.35940/ijitee.J9340.0881019.com

2). E. Taylor, P.S. Ezekiel, F.B.Deedam-Okuchaba (2019). A Model to Detect Heart Disease using Machine Learning Algorithm. International Journal of computer sciences and engineering E-ISSN:2347-2693 Vol.-7, issue-11,nov 2019https://doi.org/10.26438/ijcse/v7i11.15

3) Svetlana Ulianova, Cardiovascular Disease dataset. The dataset consists of 70 000 records of patient data, 11 features + target

.4)Seyedamin Pouriyeh, Sara Vahid, Giovanna Sannino, Giuseppe De Pietro, Hamid Arabnia, Juan Gutierrez et al. “A Comprehensive Investigation and Comparison of Machine Learning Techniques in the Domain of Heart Disease”, 22nd IEEE Symposium on Computers and Communication (ISCC 2017): Workshops - ICTS4eHealth 2017

5)Dhomse Kanchan B and Mahale Kishor M. et al. “Study of Machine Learning Algorithms for Special Disease Prediction using Principal of Component Analysis”, 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication.

6) Soni Jyoti. "Predictive data mining for medical diagnosis: An overview of heart disease prediction." International Journal of Computer Applications 17.8 (2011)

7)Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naïve Bayes algorithm. International Journal of Computer Science and Information Technologies.

8) R. Detrano, A. Janosi, W. Steinbrunn, M. Pfisterer, J.-J. Schmid,S. Sandhu, K. H. Guppy, S. Lee, and V. Froelicher, “International application of a new probability algorithm for the diagnosis of coronary artery disease,”

