VISUALIZING AND PREDDICTING HEART DISEASE WITH AN INTERACTIVE DASHBOARD

TEAM ID:PNT2022TMID44369

LITERATURE SURVEY

ABSTRACT:

Healthcare industries generate enormous amount of data, so called big data that accommodates hidden knowledge or pattern for decision making. The huge volume of data is used to make decision which is more accurate than intuition. Exploratory Data Analysis (EDA) detects mistakes, finds appropriate data, checks assumptions and determines the correlation among the explanatory variables. In the context, EDA is considered as analysing data that excludes inferences and statistical modelling. Analytics is an essential technique for any profession as it forecast the future and hidden pattern. Data analytics is considered as a cost effective technology in the recent past and it plays an essential role in healthcare which includes new research findings, emergency situations and outbreaks of disease. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analysing data. In this paper, the risk factors that causes heart disease is considered and predicted using K-means algorithm and the analysis is carried out using a publicly available data for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool. The paper discusses the pre-processing methods, classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

INTRODUCTION:

Now a day's heart disease is emerging as one of the most death-dealing diseases. As per a report published by the World Health Organization [WHO], heart disease is one of the most hazardous diseases to human which causes death all over the world from the last 20 years. Approx. 12 million people are dying every year, which makes it the biggest challenge for medical professionals to develop an early diagnosis of heart disease with better accuracy. In this paper, we have applied different machine learning algorithms and compared their classification accuracies. We have proposed a modified algorithm using logistic regression with principal component analysis for predicting heart disease with more accuracy on various attributes such as age, blood pressure, chest pain, serum cholesterol levels, heart rate, and other characteristic attributes, and patients will be classified according to varying degrees of coronary artery disease.

1) PAPER

- Using Machine Learning for Heart Disease Prediction

AUTHOR

- Dhai Eddine Salhi,Abdelkamel A Kamel Tari, Tahar Kechadi

YEAR

- 2021

METHOD AND ALGORITHM

- In this paper we carried out research on heart disease from data analytics point of view. Prediction of heart disease is a very recent field as the data is becoming available. Other researchers have approached it with different techniques and methods. We used data analytics to detect and predict disease's patients. Starting with a pre-processing phase, where we selected the most relevant features by the correlation matrix, then we applied three data analytics techniques (neural networks, SVM and KNN) on data sets of different sizes, in order to study the accuracy and stability of each of them. Found neural networks are easier to configure and obtain much good results.

ACCURACY

- 93%

2) PAPER

- Prediction of Heart Disease using Classification Algorithms .

AUTHOR

- Hlaudi Daniel Masethe Mosima Masethe

YEAR

- 2014

METHOD AND ALGORITHM

- The heart disease accounts to be the leading cause of death worldwide. It is difficult for medical practitioners to predict the heart attack as it is a complex task that requires experience and knowledge. The health sector today contains hidden information that can be important in making decisions. Data mining algorithms such as J48, Naive Bayes REPTREE, CART, and Bayes Net are applied in this research for predicting heartattacks.

ACCURACY

- 99%

3) PAPER

- Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms

AUTHOR

- Abdul Saboor,Muhammad Usman,Sikandar Ali,Ali Samad

YEAR

- 2022

METHOD AND ALGORITHM

- A great diversity comes in the field of medical sciences because of computing capabilities and improvements in techniques, especially in the identification of human heart diseases. Nowadays, it is one of the world's most dangerous human heart diseases and has very serious effects the human life. Accurate and timely identification of human heart disease can be very helpful in preventing heart failure in its early stages and will improve the patient's survival. Manual approaches for the identification of heart disease are biased and prone to interexaminer variability. In this regard, machine learning algorithms are efficient and reliable sources to detect and categorize persons suffering from heart disease and those who are healthy. According to the recommended study, we identified and predicted human heart disease using a variety of machine learning algorithms and used the heart disease dataset to evaluate its performance using different metrics for evaluation, such as sensitivity, specificity, F-measure, and classification accuracy. For this purpose, we used nine classifiers of machine learning to the final dataset before and after the hyperparameter tuning of the machine learning classifiers, such as AB, LR, ET, MNB, CART, SVM, LDA, RF, and XGB. Furthermore, we check their accuracy on the standard heart disease dataset by performing certain preprocessing, standardization of dataset, and hyperparameter tuning. Additionally, to train and validate the machine learning algorithms, we deployed the standard K-fold cross-validation technique. Finally, the experimental result indicated that the accuracy of the prediction classifiers with hyperparameter tuning improved and achieved notable results with data standardization and the hyperparameter tuning of the machine learning classifiers.

ACCURACY

- 96.72%

4) PAPER

- Comparative Analysis of Classification Approaches for Heart Disease Prediction.

AUTHOR

- s. M. Mahedy Hasan , Md Al Mamun ,Md. Palash Uddin, Dr Md Ali Hossain

YEAR

- 2018

METHOD AND ALGORITHM

- Heart disease is the one of the most common causes of death around the world nowadays. Often, the enormous amount of information is gathered to detect diseases in medical science. All of the information is not useful but vital in taking the correct decision. Thus, it is not always easy to detect the heart disease because it requires skilled knowledge or experiences about heart failure symptoms for an early prediction. Most of the medical dataset are dispersed, widespread and assorted. However, data mining is a robust technique for extracting invisible, predictive and actionable information from the extensive databases. In this paper, by using info gain feature selection technique and removing unnecessary features, different classification techniques such that KNN, Decision Tree (ID3), Gaussian Naïve Bayes, Logistic Regression and Random Forest are used on heart disease dataset for better prediction. Different performance measurement factors such as accuracy, ROC curve, precision, recall, sensitivity, specificity, and F1-score are considered to determine the performance of the classification techniques.

ACCURACY

- 92.76

5) PAPER

- A novel approach for heart disease prediction using strength scores with significant predictors.

AUTHOR

- Armin Yazdani, Kasturi Dewi Varathan,Yin Kia Chiam, Asad Waqar Malik Wan Azman Wan Ahmad.

YEAR

- 2021

METHOD AND ALGORITHM

- This paper is motivated by the gap in the literature, thus proposes an algorithm that measures the strength of the significant features that contribute to heart disease prediction. The study is aimed at predicting heart disease based on the scores of significant features using Weighted Associative Rule Mining.

ACCURACY

- 98%

6) PAPER

- Heart Diseases Detection by Machine Learning Classification Algorithms.

AUTHOR

- Ashapu Bhavani

YEAR

- 2022

METHOD AND ALGORITHM

- Most human deaths are caused by heart diseases. Such diseases cannot be efficiently detected for the lack of specialized knowledge and experience. Data science is important in healthcare sector for the role it plays in bulk data processing. Machine learning (ML) also plays a significant part in disease prediction and decision-making in medical care industry. This study reviews and evaluates the ML approaches applied in heart disease detection. The primary goal is to find mathematically effective ML algorithm to predict heart diseases more accurately. Various ML approaches including Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), T-Distributed Stochastic Neighbor Embedding (t-SNE), Naïve Bayes, and Random Forest were utilized to process heart disease dataset and extract the unknown patterns of heart disease detection. An analysis was conducted on their performance to examine the effecacy and efficiency.

ACCURACY

- 97%

CONCLUSION:

Heart stroke and vascular disease are the major cause of disability and premature death. Chest pain is the key to recognize the heart disease. In this work, the heart diseases are predicted y considering major factors with four types of chest pain.K-means clustering is one of the simplest d popular unsupervised machine learning algorithms. Here the datasets are clustered experience..