Literature survey

Date	19 September 2022
Team ID	PNT2022TMID28187
Project Name	Project - Visualizing and Predicting Heart Diseases with an Interactive Dash Board
Maximum Marks	2 Marks

1. Md. Touhidul Islam, Sanjida Reza Rafa, et al, "Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with k-Means", 2021.

Worldwide research shows that millions of lives lost per year because of heart disease. The healthcare sector produces massive volumes of data on heart disease that are sadly not used to locate secret knowledge for successful decision making. One of the most important aspects at this moment is detecting heart disease at an early stage. Researchers have applied distinct techniques to the UCI Machine Learning heart disease dataset. Many researchers have tried to apply some complex techniques to this dataset, where detailed studies are still missing. In this paper, Principal Component Analysis (PCA) has been used to reduce attributes. Apart from a Hybrid genetic algorithm (HGA) with k-means used for final clustering. We used the Hybrid Genetic Algorithm (HGA) for data clustering to avoid this problem. Our proposed method can predict early heart disease with an accuracy of 94.06%.

2. Mohini Chakarverti, Saumya Yadav, et al, "Classification Technique for Heart Disease Prediction in Data Mining", 2019.

Currently, a lot of methods have been implemented for prediction analysis. In the proposed study, clustering and classification of the input information for heart disease forecasting is executed with the help of k-means clustering algorithm and SVM (support vector machine) classification model on the basis of prediction analysis methods. The back propagation algorithm along with k-means clustering algorithm is applied for the clustering of information. These algorithms support to enhance the precision of prediction analysis. A heart disease data suite obtained from the UCI repository is used for judging the performance of proposed algorithm. This data suite comprises total 76 features. But, whole tests require a subset of 14 features. A comparison between proposed study and earlier method (using arithmetic mean) is performed in terms of certain parameters such as accuracy, error recognition rate and execution time.

3. **Narendra Mohan, Vinod Jain, et al, "Heart Disease Prediction Using Supervised Machine Learning Algorithms", 2021.**

Predicting and detecting cardiac disease has always been a difficult and time-consuming undertaking for doctors. To treat cardiac disorders, hospitals and other clinics are giving costly therapies and operations. As a result, anticipating cardiac disease in its early stages will be beneficial to people all around the world, allowing them to take required treatment before it becomes serious. Heart disease has been a major issue in recent years, with the primary causesbeing excessive alcohol use, tobacco use, and a lack of physical activity. Machine learning methods are utilized to forecast cardiac illnesses in this article. For training and testing, a data collection containing diverse human health parameters is used. Many AI&ML algorithms are used to predict cardiac disorders. The performance of the machine learning algorithm is compared after it has been implemented.

4. **Kummita Sravan Kumar Reddy, K.V. Kanimozhi et al, Novel Intelligent Model for Heart Disease Prediction using Dynamic KNN (DKNN) with improved accuracy over SVM", 2022.**

In comparison to Support Vector Machine, the major goal is to forecast the Novel Intelligent model for Heart Disease prediction using Dynamic KNN (SVM). Materials and Procedures: Two machine learning methods, Dynamic KNN (N=92) and Support Vector Machine (N=92), are used to predict heart disease. Dynamic KNN is a simple algorithm used for disease prediction. Heart disease dataset is used for disease prediction. For each group 20 samples are taken and it is divided into training and testing dataset. Result and Discussion: Accuracy of Dynamic KNN is 84.44% and Support Vector Machine is 67.21%. There exists an analytical significant difference between Dynamic KNN and SVM. Conclusion: Dynamic KNN appears to perform significantly better than Support Vector Machine for Novel Heart Disease Prediction.

5. **Akanksha Kumari, Ashok Kumar Mehta, et al, "A Novel Approach for Prediction of Heart Disease using Machine Learning Algorithms", 2021.**

Heart Disease is the leading causes of death and hospitalization in the world. With advancement of technology and the contribution of computer engineering, it is easy to detect heart disease and thus treatment is fast and effectively done. Machine learning nowadays is very popular in predicting disease in the medical field. In this paper, authors have tried to predict heart disease using seven machine learning algorithms and attempted to improve the accuracy of weak performing algorithms using ensemble methods like AdaBoost and voting ensemble method. The performance of Linear Discriminate Analysis is good among other algorithms, it's mean value is approximately 0.847 and mean absolute error

is 0.185, the false acceptance rate is lowest among all i.e.; 0.33 and the false recognition rate is 0.076, accuracy is somehow coming 80% which is less if compared with Logistic Regression.

6. **A. Lakshmanarao , A. Srisaila, "Heart Disease Prediction using Feature Selection and Ensemble Learning Techniques", 2021.**

   Cardiovascular diseases (heart-related diseases) are the reason for the deaths of 18 million people every year in the world. According to WHO,31% of the deaths worldwide are due to heart-related diseases. In this paper, we proposed a novel machine learning model for heart disease prediction. The proposed method was tested on two different datasets from Kaggle and UCI. We applied sampling techniques to the unbalanced dataset and feature selection techniques are used to find the best features. Later several classifier models were applied and achieved good accuracy with ensemble classifier. The experimentations on two datasets shown that the proposed model is effective for heart disease prediction. Python was used for all implementations. Experimental analysis resulted in an accuracy of 85.71% for UCI Heart Disease dataset and 87.30% for Cardiovascular Disease dataset.

7. **Vaibhav Gupta, Vaibhav Aggarwal, et al, "Visualization and Predictionof Heart Diseases Using Data Science Framework", 2021.**

   Heart is one the most vital organ in the human body. When we talk about heart diseases, we can have multiple conditions where heart is not working the way it should be like blockage in blood vessels. According to many researches that have been conducted through a period of time have found out that heart failure and heart disease has been the cruel cause of death in human beings. What aggravates this situation is that most of these diseases are being diagnosed at later stages at which it is very difficult to control. But if somehow, we can diagnose these diseases at its early stage, then we can surely cure the disease. The main aim of this paper is to use various classification algorithms of data science framework to somehow detect the chances of having a heart disease. Also, the main aim of this research paper is to find out the most efficient classification algorithm that can help us to detect heart diseases at early stage. this algorithm can be used on heart records of the patient or by using it on classification reports. This research was conducted and tested upon various algorithms to test its accuracy like Logistic Regression, Random Forest, Vector Support and XG-Boost. After applying these algorithms of prediction model has been developed.

8. **Sakshi Bhoyar, Nikki Wagholikar, et al, "Real-time Heart Disease Prediction System using Multilayer Perceptron", 2021.**

   Stroke, Heart Failure, Arrhythmia, and myocardial infarction are the most common cardiovascular diseases which record high mortality rates around the world. Heart defects are not detected in the early stages due to the impractical costs of the tests available. Thus, a fast, real-time, and reliable system that predicts the chances of a patient having heart disease in an optimized manner is required.In this research, a Neural Networks model using a Multilayer Perceptron (MLP) is proposed for the prediction system. Experimental analysis resulted in an accuracy of 85.71% for UCI Heart Disease dataset and 87.30% for Cardiovascular Disease dataset. When compared to previous research the increase in accuracy was approximately 12-13%. A simple web application tool is also developed using python programming to test the prediction system. This research works towards making a comprehensible tool for medical professionals as well as common people.

9. **D.P. Yadav, Prabhav Saini, et al, "Feature Optimization Based Heart Disease Prediction using Machine Learning", 2021.**

   Heart disease is a spontaneous, treacherous, and fatal disease. It is a group of several states that result in abnormal functioning of the heart. Based on the several pathology test report heart disease is identified by a doctor. The manual heart disease prediction is time consuming and error prone. Therefore, in the present study an automated system based on the performance analysis of several machine learning techniques has been developed. First, the well-known machine learning algorithm Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes and Random Forest applied on the dataset for the prediction of heart disease. To avoid bias performance 3-fold cross validation is applied. The highest average accuracy of 87.78 % is obtained by the Naïve Bayes. The performance of the model is acceptable. Further, we have applied genetic algorithm on the dataset to optimize the features. After, optimization the highest average accuracy of 96% is achieved by the naïve Base.

10. **Lakshmi Prasad Koyi, Tejaswi Borra "A Research Survey on State of the art Heart Disease Prediction Systems", 2021.**

    Disease prediction systems are the better alternatives, to avoid the human errors in disease diagnosis and also assist in disease prevention with early detections. Highdemand in preventing the rapidly increasing heart disease death tolls expanded the horizons of the former research scholars for introducing the intelligent heart disease prediction systems. Prediction of heart disease from patient's health record attributes is, a proven multi-dimensional decision-making system, which merely depends on mining attribute correlations too. Patient

Health Records (PHR's) with structured categorical data and unstructured text/image data are the major input resources for heart disease prediction. Heart disease dataset preparation, prediction system's process flow design, process execution and results evaluation are the most common life cycle modules of any heart disease prediction system. Although many former research were introduced various heart disease prediction models, but they are still suffering from some common set of problems. Input dataset attributes modeling, attributerisk factor calculation, correlations mining; threshold determination and achieving the high accuracy in disease prediction are the major limitations of the existing heart disease prediction systems. As part of my research on designing intelligent heart disease prediction models, several research papers are analyzed and narrated that knowledge in a proper manner with detailed description. The main objective of this study is to represent the current scenario of heart disease prediction systems and their associated modules in brief.