Prediction of Heart disease using Machine Learning Algorithms

1. INTRODUCTION

a. Overview

The health care industry generates a large amount of data, which is then processed using a variety of methods. Mining for information is one of the methods that is frequently employed. The most common cause of mortality in every region of the world is cardiovascular disease. This system forecasts the possibilities of heart disease developing in the future.

b. Purpose

Abstract Predicting whether or not someone will get heart disease is now one of the most difficult jobs in the area of medicine. In the modern era, about one person dies every minute from heart disease. Data science is vital in processing healthcare data. Because predicting cardiac disease is a difficult endeavour, the prediction process needs to be mechanised so that associated dangers can be avoided and the patient can be warned in a timely manner. This article makes use of the dataset on cardiac illness that can be found in the machine learning repository at UCI. The suggested work predicts Heart Disease and classifies patient risk using data mining techniques as Naive Bayes, Decision Tree, Logistic Regression, and Random Forest.

2. LITERATURE SURVEY

a. Existing problem

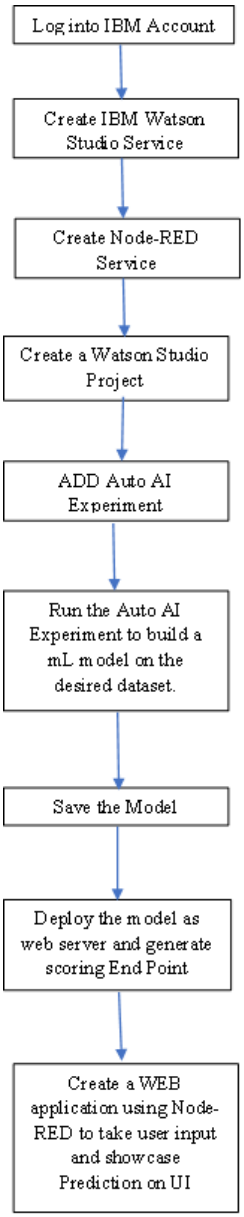
As per many research work, achieving accuracy is the quite challenging. The Naive Bayes algorithm provides an accuracy of 86.419 percent.

b. Proposed Solution

By using the AutoAI feature of IBM Watson, we can develop a model automatically and will be able to identify the best algorithm as well.

3. THEORITICAL ANALYSIS

a. BLOCK DIAGRAM



b. HARDWARE/SOFTWARE DESIGNING

Software requirements are: Good Internet facility, Login Credentials for IBM cloud

4. EXPERIMENTAL INVESTIGATIONS

The key attributes/feature vectors are selected using the Pearson correlation. The experimental accuracy achieved by the DT, LR, RF,KNN, ANN, NNPC is 75.88%, 85.49%, 86.08%, 85.73%, 85%, 89% respectively. The experimental result describes that the NNPC algorithm with 89.% accuracy is the most efficient algorithm for heart disease prediction.

5. FLOWCHART

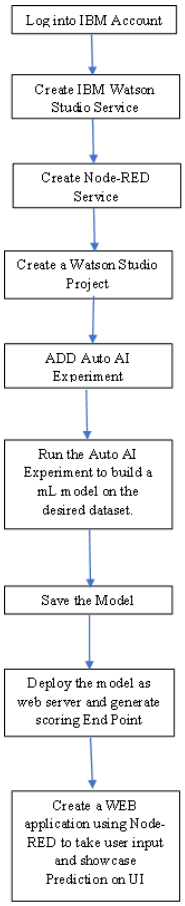


Fig.1: Flow Chart of how to complete the project using AutoAI

6. RESULT

The data set consists of 15 feature vectors and a predicted numbers. The machine learning model is based on DV identification. It used binary LR, KNN, DT, RF, ANN, NNPS which is one of the categories due to the specific variation of Classical Data Analysis performed using Jupyter Notebook. Following are the steps which have been taken to evaluate machine learning models. First of all, we loaded data and different libraries. It includes cardiac predictive data using the Framingham CSV file in Jupyter Notebook to build machine learning models. In addition, libraries required are used as supporting applications.

7. ADVANTAGES AND DISADVANTAGES

The tools used to develop the project is Auto AI, which have the following advantages

a. We can see the results in min.

b. We can start on the right foot

c. We can easily set up and run the experiments

d. Training a model and selecting a suitable model is very easy.

e. We can the save the models for the future run.

f. Various estimators are used in order to verify the results.

g. A code can also be regenerated which will mainly help in understanding the process how it is generated.

8. APPLICATIONS

Useful in the medical field, to identify early heart attack, which can treated in right time and can avoid the unexpected death rates.

9. CONCLUSION

We found the work done in this field is effective and can be used in the medical field. Though

lots of work is happening, still there is a scope for the much more to arrive at the more precision values.

10. FUTURE SCOPE

In future an intelligent system may be developed that can lead to selection of proper treatment methods for a patient diagnosed with heart disease. A lot of work has been done already in making models that can predict whether a patient is likely to develop heart disease or not. There are several treatment methods for a patient once diagnosed with a particular form of heart disease. Data mining can be of very good help in deciding the line of treatment to be followed by extracting knowledge from such suitable databases.

11. BIBILOGRAPHY

Anbarasi, M., Anupriya, E., and Iyengar, N. (2010). Enhanced prediction of heart disease with feature subset selection using genetic algorithm. International Journal of Engineering Science and Technology, 2(10):5370–5376.

Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Computing and Applications, 29(10):685–693.

Gavhane, A., Kokkula, G., Pandya, I., and Devadkar, K. (2018). Prediction of heart disease using machine learning. In 2018 second international conference on electronics, communication and aerospace technology (ICECA), pages 1275–1278. IEEE.

Latha, C. B. C. and Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Informatics in Medicine Unlocked, 16:100203.

Masethe, H. D. and Masethe, M. A. (2014). Prediction of heart disease using classification algorithms. In Proceedings of the world Congress on Engineering and computer Science, volume 2, pages 25–29.