Project Report

1 INTRODUCTION

1.1 Overview

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide.

Heart failure is a common event caused by CVDs and predicting mortality using some simple mechanism is required.

1.2 Purpose

Create a ML based model to detect/predict heart diseases using some ML based approach

2 LITERATURE SURVEY

2.1 Existing approaches

Annepu et.al [1] proposed a heart disease prediction framework based on RF algorithm in machine learning using Python. They used the Cleveland heart disease datasets obtained from the UCI machine learning repository for the algorithm training and testing. This sample originally contains 303 instances with 76 features but after preprocessing and manual attribute selection of features, only 9 features were used. 75% of the sample was used for algorithm training while 25% was used for testing. A graphical user interface (GUI) was developed using Visual Studio Code for visualization of the experiments. The RF classifier was employed for the classification, where an accuracy of 97.56% was achieved.

In [2], authors proposed a heart disease prediction based on machine learning techniques using NB and DT algorithms in Python. The datasets used for training and testing of the model were obtained from the Kaggle website, which contain 13 heart disease features. Another dataset from the UCI machine learning repository was used for the simulation. The proposed model was implemented on the Scipy environment. Form their experiments, results showed that DT algorithm performed better than the NB in the prediction of heart diseases. Their study had a lot of shortcomings, which include unspecified datasets, unavailability of real experiments, imprecise results, and improper feature selection approach.

Nashif et al. [3] proposed a tentative design of a cloudbased heart disease prediction system using machine learning techniques. Two of the UCI datasets: Cleveland heart disease data consisting of 303 instances with 14 features and VA Long Beach data consisting of 270 instances with also 14 features were merged together making a bigger dataset. Five machine learning algorithms, including MLP,

LR, NB, RF, and SVM in the Java-based open access platform (WEKA) were applied in the classification and prediction processes. Of the five algorithms, SVM appeared the best classifier with a classification accuracy of 97.53%.

Patel et. al [4] compared the performances of J48, Logistic Model Tree (LMT), and Random Forest (RF) algorithms in WEKA for heart disease prediction task. The Cleveland datasets from the UCI, which commonly consist 303 instances and 76 features were used for the training and testing using the 10-fold cross validation method. The evaluation was based on accuracy, sensitivity, and specificity. Experimentation was carried out on Core i3 with 2.4GHz CPU and 4GB RAM. Results showed that J48 appeared with the highest classification accuracy of 56.76% followed by LMT at 55.77%, then RF.

In [5], authors proposed an intelligent system framework for heart disease prediction using NB classifier, which was implemented on java platform. The study was carried out in comparison with DT algorithm prediction performance. In their implemented system, patients were to enroll their required information which would be stored in the system database, and the classification would be done automatically during enrollment. Upon entering the information, patients would be classified as either heart disease or normal. The information is viewed by a medical professional.

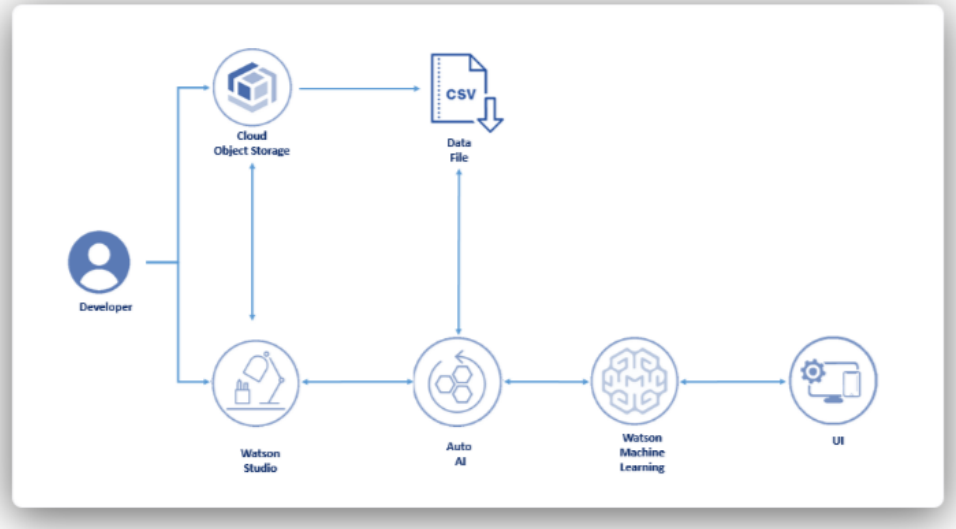
2.2 Proposed solution

In this project, a model using Auto AI and a web application was created to showcase the prediction of heart failure.

THEORITICAL ANALYSIS

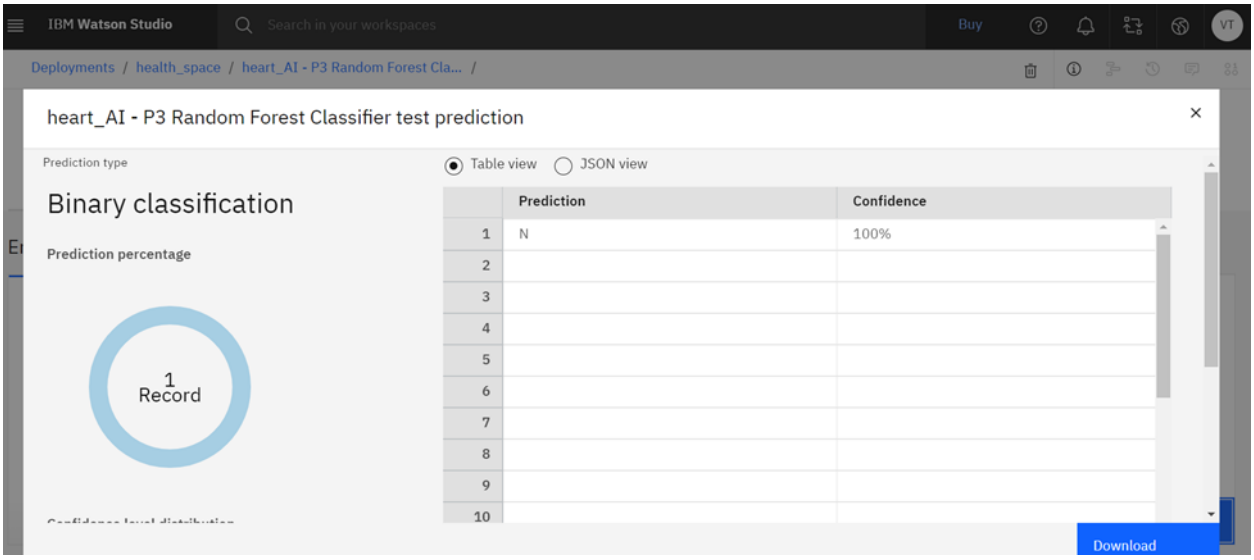
3.1 Architecture

Technical Architecture:



The above figure shows the complete block diagram presenting the steps involved in successfully deployment of the ML model using IBM - auto AI - node red

4 RESULT



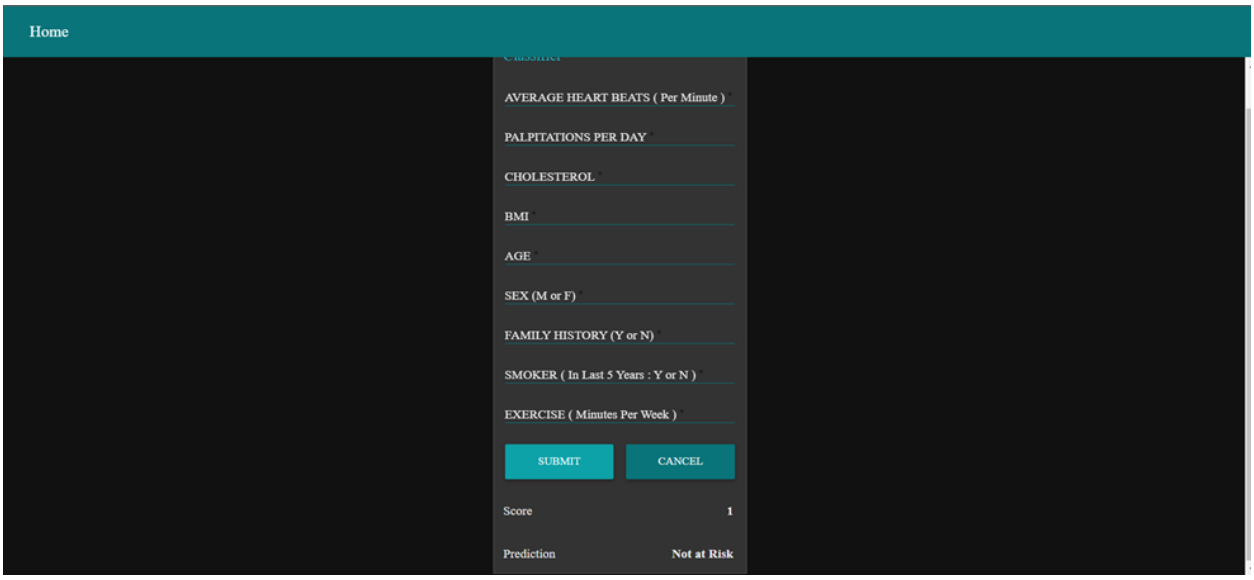


FIg 1. presents the performance of the built model for some information of a random person. It was observed that the model is predicting as no risk case with a confidence level of 100%

Fig 2. presents the web application of the model deployed using node-red application. This node-red application uses the AI model deployed in the ibm cloud for predicting CVDs

5 ADVANTAGES & DISADVANTAGES

Advantages:

IBM cloud, watson studio, auto AI, node -red are integrated with a single platform and this helps in easy deployement of model. This type of integration helps in robust building of end-to-end solution with in a short peiod of time.

disadvantages:

This ibm model using auto AI is not tested on new datasets.

6 APPLICATIONS

Cardiovascular disease detection

continuous monitoring of heart patients

7 CONCLUSION

In this work, we had implemented a ML model using IBM watson studio and observed that random forest classifer is performing best with a accuracy of 87%. The model is deployed using web application and tested its accuracy for various combinations of input features.

8 FUTURE SCOPE

The presented work can be extended to real world scenarios by adding more features and testing the robustnes of the model for diverse population. Also, image

processing based cardiovascular disease detection can be implemented using DL/ML.

BIBILOGRAPHY

[1] Annepu, D., & Gowtham, G. (2019). Cardiovascular disease prediction using machine learning techniques. International Research Journal of Engineering and Technology, 6 (4), 3963-3971.

[2] Sridhar, A., & Kapardhi, A. (2018). Predicting heart disease using machine learning algorithm. International Research Journal of Engineering and technology, 6 (4), 36-38

[3] Nashif, S., Raiban, M., Islam, M., & Imam, M. H. (2018). Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system. World Journal of Engineering and Technology, 6, 854-873.

[4] Patel, J., Upadhyay, T., & Patel, S. (2016). Heart disease prediction using machine learning and data mining techniques. IJCSC, 7 (1), 129-137.

[5] Ritesh, T., Gauri, B., Ashwini, D., & Priyanka, S. (2016). Heart attack prediction system using data mining. International Journal of Innovative Research in Computer and Communication Engineering, 4 (8), 15582-15585.

.