

Project Title: Heart Rate Disease Prediction using Azure

Team Members:

Varun Simha Reddy Thippireddy
Sri Vidya Bachina
Sai Vignesh Rayal Rolla

Goals

- Background

Nowadays, heart disease become a thing of concern for sudden death in young generations. It is estimated that approximately 1.7 million people died in the previous year due to heart attacks. There has been an exponential increment in heart disease cases over the past few years. The prediction of heart disease is a critical process that needs to be detected at an early stage. Several distinct symptoms are there in heart disease which makes it very difficult to predict. The heart is a major organ of the human body which plays a critical role in pumping blood throughout the body with the help of blood vessels. Heart disease is one type of disease that affects the heart, and the prediction of the heart thus becomes very important to minimize death. The primary goal of this research work is to determine the best machine learning model which can effectively predict heart disease along with various key factors. The overall deployment of the prediction model will be performed using the Azure machine-learning model. The research work should be performed in such a way that successfully generates the required output for this project work. Moreover, a suitable heart disease dataset will be taken into consideration for performing the desired task effectively. The dataset should include some major features such as age, sex, various types of chest pain, cholesterol, resting ECG, old peak, and heart disease. The deployment of the Microsoft Azure software needs to be performed in such a way that can effectively predict heart disease in the human body. Moreover, in this project work, the utilization of several machine learning algorithms such as Decision Forest, Logistic Regression, and SVM, Random Forest, KNN is going to be done successfully.

- Significance

Heart disease has become one of the major causes of human loss in recent times all over the world. Moreover, it is observed that nowadays heart disease has appeared in middle age groups. Therefore, it is very important to monitor health by detecting deviations in major features that are injurious to health. The primary focus of this project work is to identify the possible cause of heart attacks using machine learning algorithms (Mohan *et al.*, 2019). The observation of the patient's condition should be based on some critical features such as the age of the patient, chest

pain type, cholesterol, exercising angina, and several other features. Therefore, it can be stated that the overall implementation of machine learning algorithms can be beneficial for detecting heart disease. However, there are several issues associated with heart disease prediction that should be mitigated properly to obtain the desired output in this project work. The exponential growth in cardiovascular disease in recent times causes huge human losses. Therefore, it becomes very essential for developing a suitable machine-learning model which can effectively predict heart disease. The overall utilization of all the required machine-learning models should be performed in such a way that can detect heart disease in the human body in an efficient manner. However, it is observed that there are several issues associated with heart disease prediction which need to be mitigated properly so that better output can be generated.

- Objectives

- ⇒ To determine the best-suited machine learning model for predicting heart disease
- ⇒ To implement Microsoft Azure cloud services software for developing the heart disease prediction model
- ⇒ To understand the critical features of the machine learning models
- ⇒ To identify the deviation in major features that can affect the health
- ⇒ To implement various machine learning algorithms such as Logistic regression as well as SVM along with decision forest.

- Features

The deployment of the heart disease prediction model should be performed with the help of Microsoft Azure software. The Azure machine learning model is a part of Microsoft Azure cloud services which contain several features which can be utilized in this project work to manage various machine learning projects. The management of the machine learning project life cycle can be effectively obtained with the help of this software. The data training and test requires to be performed with the help of a machine learning technique. The prediction model can be generated effectively after completing the data training successfully (Yahaya *et al.*, 2020). The deployment of the prediction model should be performed after successfully completing the development process. Moreover, the machine-learning model should contain several critical features which can successfully predict heart diseases. The machine learning model contains several features such as data processing, feature engineering, selection of proper algorithms, diverse algorithms, data modeling, and various critical features which can be utilized for predicting the heart disease successfully. Thus, the overall implementation of the machine learning model can be beneficial for developing a suitable model which can effectively predict heart disease (Fitriyani *et al.*, 2020). All the required features of the machine learning model can be utilized properly for the purpose of generating the required objectives.

- Dataset

We are using the (heart.csv) a data file considered from Kaggle(<https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>)

⇒ Loading the data to Azure portal to implement the different Algorithm's.

- Implementation:

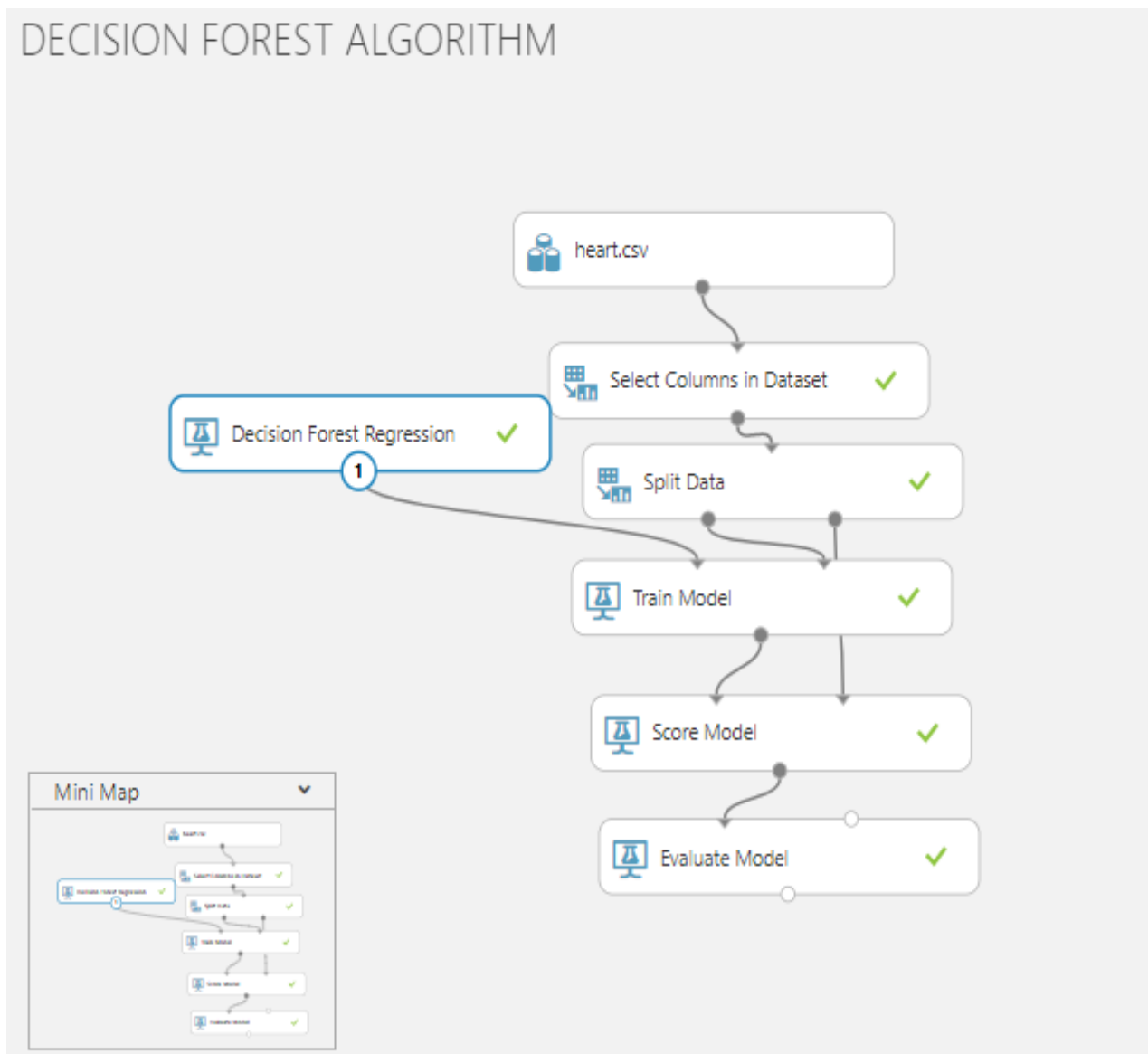


Figure 2: Decision Forest model

(Source: Self-developed)

The model has been built for decision forest and after importing of dataset the selection of column has been done and the output has been chosen as the target column. Then after that, the score model and the evaluation model have been built and after that, the decision tree has been perfectly analyzed here.

DECISION FOREST ALGORITHM > Decision Forest Regression > Untrained model

Gemini Decision Forest Regressor

Settings

Setting	Value
Ensemble Element Count	8
Max Depth	32
Random Split Count	128
Min Leaf Sample Count	1
Class Count	1
Resampling Method	Bagging
Random Number Seed	5
Allow Unknown Levels	True

Figure 3: Untrained model of decision tree

(Source: Self-developed)

This is the untrained model of the heart dataset is shown with the effective values.

(Source: Self-developed)

This has been used as the trained model, and all the formation of the decision tree with the correct model has been mentioned here in this graph visualization.

- Preliminary Result:

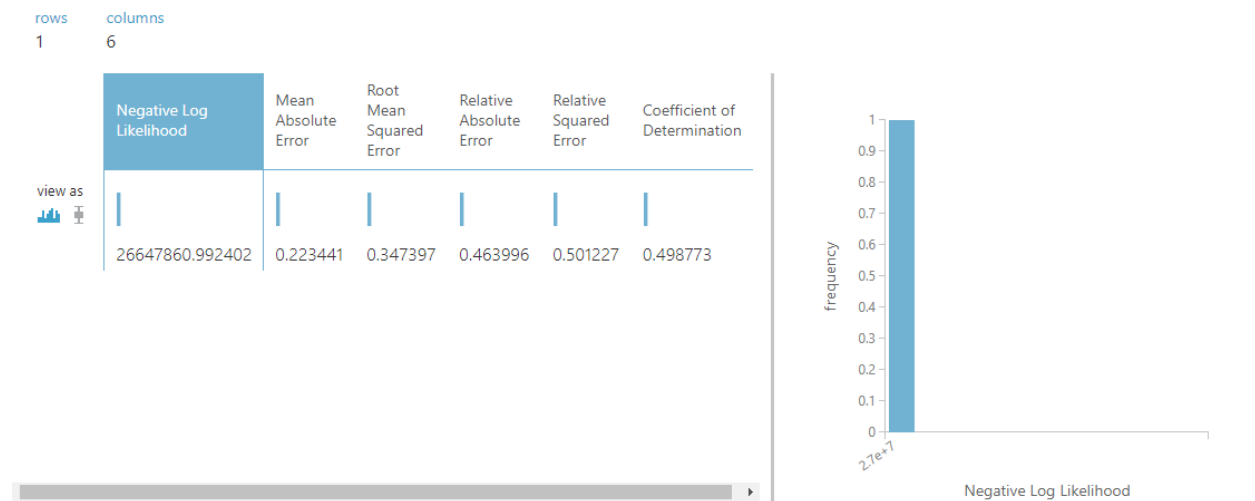


Figure 5: Evaluation result of the decision tree

(Source: Self-developed)

- Code
- Responsibilities

Subject	Assignment	Status	Work Assigned to
Project Planning	Determine project potentiality	Completed	Varun simha reddy, Sri vidya, Sai Vignesh
Project planning	Extracting the data and its Features	Completed	Sri Vidya, Sai Vignesh
Choosing Algorithms	Exploring different Algorithms for the model	Completed	Varun Simha Reddy
Algorithm Implementations	Implementing the Decision Forest	Completed	Sri Vidya
Algorithm Implementations	Implementing the SVM	In progress	Sai Vignesh Rayal
Algorithm Implementations	Implementing the Logistic Regression	In progress	Varun Simha Reddy
Model Analysis Exploration	Testing & exploring the results and implementing different algorithms	In progress	Sri Vidya, Sai Vignesh
Documentations	Documentations	In progress	Varun Simha Reddy

Reference:

1. Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 63, pp.208-222.
2. Ayon, S.I., Islam, M.M. and Hossain, M.R., 2022. Coronary artery heart disease prediction: a comparative study of computational intelligence techniques. *IETE Journal of Research*, 68(4), pp.2488-2507.
3. Fitriyani, N.L., Syafrudin, M., Alfian, G. and Rhee, J., 2020. HDPM: an effective heart disease prediction model for a clinical decision support system. *IEEE Access*, 8, pp.133034-133050.
4. Khan, M.A., 2020. An IoT framework for heart disease prediction based on MDCNN classifier. *IEEE Access*, 8, pp.34717-34727.
5. Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, 7, pp.81542-81554.
6. Rajdhan, A., Agarwal, A., Sai, M., Ravi, D. and Ghuli, P., 2020. Heart disease prediction using machine learning. *International Journal of Research and Technology*, 9(04), pp.659-662.
7. Repaka, A.N., Ravikanti, S.D. and Franklin, R.G., 2019, April. Design and implementing heart disease prediction using naives Bayesian. In *2019 3rd International conference on trends in electronics and informatics (ICOEI)* (pp. 292-297). IEEE.
8. Shah, D., Patel, S. and Bharti, S.K., 2020. Heart disease prediction using machine learning techniques. *SN Computer Science*, 1(6), pp.1-6.
9. Yahaya, L., Oye, N.D. and Garba, E.J., 2020. A comprehensive review on heart disease prediction using data mining and machine learning techniques. *American Journal of Artificial Intelligence*, 4(1), pp.20-29.