IBM project

Visualizing and Predicting Heart Diseases with an Interactive DashBoard

Bonafide record of work done by

SARUBALA P	(19EC027)
SIVASANGARI B	(19EC030)
MONISHA C	(20LEC043)
VANISHREE M	(20LEC044)

19Z039 – PROFESSIONAL READINESS FOR INNOVATION , EMPLOYABILITY AND ENTREPRENEURSHIP

FACULTY MENTOR : KARTHIKEYANI A

Problem Statement :

Visualizing and Predicting Heart Diseases with an Interactive Dashboard.

Abstract :

The machine learning models for heart disease prediction are described in this work. Medical professionals can assist patients by identifying heart illness before it manifests. The leading cause of death nowadays is heart disease. In the clinical setting, heart disease prediction is challenging. But today, a number of methods for predicting heart illness have been developed. These prediction is based on the previous or current medical record's provided by the patient or doctor .The analysis of data is on interactive dashboard where user can enter medical record's and view the predicted result. In this overview paper, methods for predicting heart disease were discussed.

Introduction :

Heart disease defines a condition that affects the blood vessels or the heart. Smoking, high blood pressure, high cholesterol, a poor diet, insufficient exercise, and obesity may all raise one's chance of developing various cardiac problems .Blood arteries that are restricted or obstructed as a result of heart disease can cause a number of diseases that might cause chest pain, a heart attack, or a stroke. Aside from those that affect the muscle, valves, or rhythm of your heart, there are other cardiac disorders that are regarded as types of heart disease.

Literature survey:

To predict the heart disease we can use different type of ML models

Logistic Regression :

The method of modeling the probability of a discrete result given an input variable is known as logistic regression. The most frequent logistic regression models have a binary outcome, which might be true or false, and so forth. It predicts the output class of the given features and gives a probabilistic value which lies between 0 and 1 as a result.

$$log\left(\frac{}{1-P}\right) = a + bX \qquad (2)$$

$$P = \frac{e^{a+bX}}{1+e^{a+bX}} \qquad (3)$$

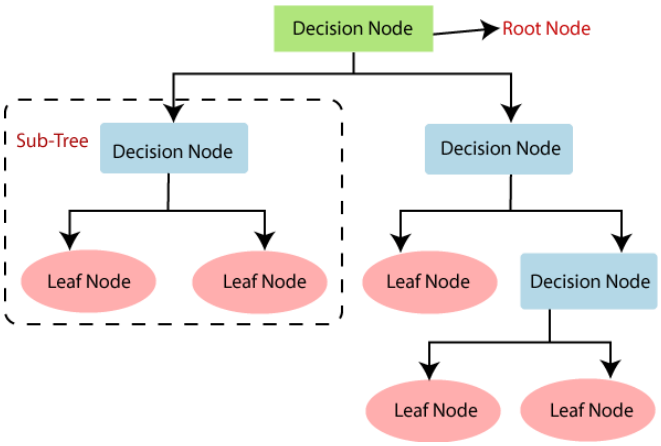
where P is the probability that a review belongs to a class

a and b are regression coefficients

X is the feature vector

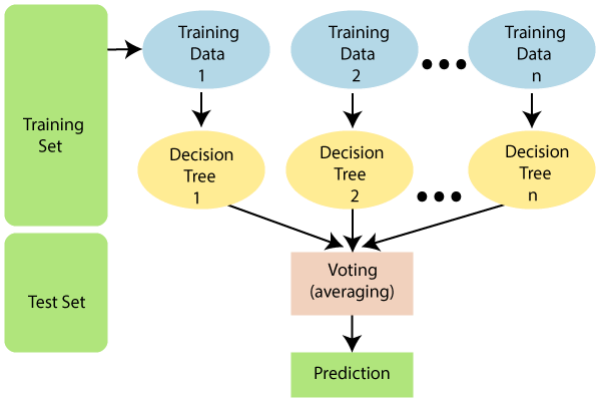
Decision Tree :

A decision tree is a tree-structured classifier that consists of a root node, a decision node, and a leaf node, where nodes represent dataset attributes, branches illustrate decision rules, and the consequence of that decision node is a leaf node.



Random Forest:

Random Forest is a classifier that incorporates multiple decision trees on diverse subsets of the given dataset and takes the mean to enhance the predictive accuracy of that dataset.



Recurrent Neural network :

With today's increasing computational power, deep learning has been used to build many complex neural networks, such as convolutional neural networks, recurrent neural networks , and depth neural networks . RNN is suitable for time serious prediction problem. RNN consist of an input layer, a hidden layer, and output layer. The result of the hidden layer is related to the input of the current layer and the output of the previous layer.

Data description :

Dataset consists of public medical records. These records include tobacco, type A (behaviour exicit by a person), family record, alcohol consumption, Adiposity(BMI), sbp (Systolic blood pressure), Age, LDL (low-density lipoprotein).

Fields	**Values**
LDL	100 - 189 mg/dl
Adiposity	Below 18.5(under weight), above 24.9 (over weight)
Alcohol Consumption	0-150 (self explanatory)
Age	1-100 (change according to people

	choice)
sbp	120 - 150 mm Hg (it may vary)
tobacco	In kg
Type A	1-100
Family record	0 or 1

Reference :

1 .Ashir Javeed, Shijie Zhou et al. (2017) designed “An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection”.

2. Bo Jin, Chao Che et al. (2018) proposed a “Predicting the Risk of Heart Failure With EHR Sequential Data Modeling”.

3. Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava, “Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques”

4. L. Ali, A. Niamat, J. A. Khan, N. A. Golilarz, X. Xingzhong, A. Noor, R. Nour, and S. A. C. Bukhari, “An optimized stacked support vector machines based expert system for the effective prediction of heart failure,”

5. A. K. Dwivedi, “Performance evaluation of different machine learning techniques for prediction of heart disease,”

6. M.Satish, D Sridhar, “Prediction of Heart Disease in Data Mining Technique”.