MACHINE LEARNING FRAMEWORK FOR HEART FAILURE PROGNOSTICS

Catherine Prashanthini Andra Thomas Balraj
Illinois Institute of Technology
candrathomasbalraj@hawk.iit.edu

PROBLEM DESCRIPTION

The healthcare industry generates a huge amount of data on a daily basis from Electronic Health Records (EHR) , In-Hospital Records (IHR) including vitals of the patient. The healthcare sector is “data rich” but “knowledge poor” and this is because not all the data are utilized for discovering patterns in occurrence of diseases and for effective decision making. These can also be used to analyse the quality of care provided by the hospitals. Efficient methodologies and tools can be used to extract such information and run analysis of them.

SOURCE:https://journalbitcoin.com/healthcare-analytics-market-2019-global-industry-analysis-type-applications-players-size-share-growth-trends-and-forecast-to-2025

The World Health Organization has estimated 12 million deaths occur worldwide, every year due to Heart diseases. Most cardiovascular diseases can be prevented by addressing behavioral risk factors such as tobacco use, unhealthy diet and obesity, physical inactivity and harmful use of alcohol. People with cardiovascular disease or

who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, etc.) need early detection and management. The early prognosis of cardiovascular diseases can aid in making decisions on lifestyle changes in high risk patients and in turn reduce the complications.

The Framingham heart study dataset from Kaggle is used to predict the Coronary Heart Disease (CHD) with the features such as gender, age, Smoking, BP Medications, Prevalent stroke, Prevalent Hypertension, Diabetes, Total cholesterol, Systolic BP, Diastolic BP, BMI, Heart rate, Glucose. These features are used to create a prognostic model that predicts the occurrence of heart failure using supervised machine learning algorithms.

LITERATURE SURVEY AND PROJECT NOVELTY

Numerous studies have been done that have focused on diagnosis of heart disease. Different data mining techniques and multiple algorithms have been used for diagnosis and achieved various probabilities for different methods.

The objective of this project is to develop an ensemble machine learning algorithm to predict heart failure and to create a web framework for getting the patient data and predicting the result in real time. Recommendation can also be included if the patient has high risk of heart attack such as insurance policies suitable for the patient.

PRELIMINARY PLAN (MILESTONES)

Collection and Preprocessing of Data - 15th March

Feature engineering and primary modelling - 1st April

Secondary modelling, prediction and design of web framework - 15th April

Recommendation system - Final presentation.

REFERENCES

[1] J. Thomas and R Theresa Princy (2016) “Human heart disease prediction system using data mining techniques” - 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT).

[2] Sellappan Palaniyappan ; Rafiah Awang (2008), “Intelligent heart disease prediction system using data mining techniques“ - 2008 IEEE/ACS International Conference on Computer Systems and Applications.

[3] Abhishek Rairikar, Vikas Sabale, Vedant Kulkarni, Harshavardhan Kale and Anuradha Lamgunde (2017), “Heart disease prediction using data mining techniques“ - 2017 International Conference on Intelligent Computing and Control (I2C2).

[4] K. Sudhakar, M. Manimekalai, (2014) "Study of Heart Disease Prediction using Data Mining", *International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, no. 1*