V.S.B.ENGINEERING COLLEGE, KARUR

Department of Electronics and Communication

Engineering

IBM NALAIYA THIRAN

TITLE   : Visualizing and predicting heart

diseases with an interactive dashboard.

DOMAIN NAME   :DataAnalytics

LEADER NAME   : Abirami T

TEAM MEMBER

NAME   **:**A Afiya fargath

                       Jenifer G

                       Jothika M G

MENTOR NAME   : Janani S

ABSTRACT

   Healthcare industries generate enormous amount of data, so called big data that accommodates hidden knowledge or pattern for decision making. The huge volume of data is used to make decision which is more accurate than intuition. Exploratory Data Analysis (EDA) detects mistakes, finds appropriate data, checks assumptions and determines the correlation among the explanatory variables. In the context, EDA is considered as analyzing data that excludes inferences and statistical modeling. Analytics is an essential technique for any profession as it forecast the future and hidden pattern. Data analytics is considered as a cost effective technology in the recent past and it plays an essential role in healthcare which includes new research findings, emergency situations and outbreaks of disease. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analyzing data.

In this paper, the risk factors that causes heart disease is considered and predicted using K-means algorithm and the analysis is carried out using a publicly available data for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool. The paper discusses the pre-processing methods, classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

Introduction

Exploratory Data Analysis (EDA) is a method to analyze data using advanced techniques to expose hidden structure, enhances the insight into a given dataset, identifies the anomalies and builds parsimonious models to test the underlying assumptions. Predicting and diagnosing heart disease is the biggest challenge in the medical industry and it is based on factors like physical examination, symptoms and signs of the patient [1-3]. Factors which influence heart diseases are cholesterol level of the body, smoking habit, and obesity, family history of diseases, blood pressure and working environment. Machine learning algorithms play a vital and accurate role in predicting heart disease .The advancement of technologies allows machine language to pair with big data tools to handle unstructured and exponentially growing data.

The invasive techniques are implemented to diagnose heart diseases based on medical history, symptom analysis report by experts, and physical laboratory report. Heart disease can be predicted based on various symptoms such as age, gender, pulse rate etc. Data analysis in healthcare assists in predicting diseases, improving diagnosis, analyzing symptoms, providing appropriate medicines, improving the quality of care, minimizing cost, extending the life span and reduces the death rate of heart patients. ECG (Electro Cardio Gram) helps in screening irregular heart beat and stroke with the embedded sensors by resting it on a chest in order to track the patient's heart beat. Heart disease prediction is being done with the detailed clinical data that could assist experts to make decision. Human life is highly dependent on proper functioning of blood vessels in the heart. The improper blood circulation causes heart inactiveness, kidney failure, imbalanced condition of brain, and even immediate death also. Some of the risk factors that can cause heart diseases are obesity, smoking, diabetes, blood pressure, cholesterol, lack of physical activities and unhealthy diet. Acute Myocardial Infarction (AMI) is the cardiovascular disease that happens due to interruption.

Literature survey

Lean Analytics: Use Data to Build a Better Startup Faster, by Alistair Croll and Benjamin Yoskovitz, However, due to its vast application, predictive analytics should not concern only business professionals. Most people are aware that companies collect our GPS locale, text messages, credit card purchases, social media posts, Google search history, etc., and this book will give you an insight into their data collecting procedures and the reasons behind them.

Eric Siegel's data analytics book is an eye-opening read for anyone who wants to learn what predictive technologies are, and how they can be deployed across a wide range of disciplines. It is not a manual, so a data scientist looking for instructions would be disappointed. Although there is some discussion of algorithms including linear regression or decision trees, it's easy to understand even for a layman.

Siegel's research makes it clear that predictive analytics is not a sneaky procedure used by companies to sell more, but a significant leap in technology that, by predicting human behavior, can help combat financial risk, improve health care, reduce spam, toughen crime-fighting, and yes, boost sales. It was lately revised and updated in January 2016.

REFERENCES

1. Adams, M.N.: Perspectives on Data Mining. International Journal of Market Research
52(1), 11–19 (2010)
2. Asur, S., Huberman, B.A.: Predicting the Future with Social Media. In: ACM International
Conference on Web Intelligence and Intelligent Agent Technology, vol. 1, pp. 492–499
(2010)
3. Bakshi, K.: Considerations for Big Data: Architecture and Approaches. In: Proceedings of
the IEEE Aerospace Conference, pp. 1–7 (2012)
4. Cebr: Data equity, Unlocking the value of big data. in: SAS Reports, pp. 1–44 (2012) .