Prediction of Heart Disease using Multiple Linear Regression Model

1. Problem Definition:

In the field of Medical Science, Heart disease predation is one of the growing areas for prediction. Huge amount of patient related data is maintained on daily basis. The stored data can be used as a source of predicting the chance of future diseases that makes the data mining techniques to play a central role for the extraction of knowledge and prediction. In this paper, Multiple Linear Regression Analysis has been performed to predict the chance of heart disease.

2. Motivation:

- According to the American Heart association, heart disease kills one person every 40 seconds, thus to identify it before any major issues happen and to get eradicate from this can save many life's.
- Helpful to doctors to know about heart disease well before.
- This is a major necessity for improving the country's healthcare sector.
- This analysis can also determine role of different symptoms leading to heart disease which can be used to take precautions.
- Identifying heart disease is quite challenging task as a number of medical reports and test are needed to be carried out also this is very costly procedure which may be not affordable for some people.
- Heart disease is kind of mysterious disease which can be considered to be encountered by false assumption which may mislead person.

3. Report works of paper:

- In this research paper Author K.Polaraju and D.Durga Prasad describe the method for predicting Heart disease using Data mining technique.
- Here authors use Heart disease dataset by UCI to train machine, which include following attributes.

    - Sex
    - Fasting Blood Sugar
    - Exang
    - No. of major vessels
    - Chest Pain Type
    - Restecg
    - Slope
    - Thal (maximum heart rate

achieved)
- Trest Blood Pressure
- Thalach
- Serum Cholesterol
- ST depression and Age

- Multi linear regression technique is used on this dataset for prediction.
- Through the paper author started with giving an abstract about heart disease and related details followed with introduction which includes different data mining techniques that can be applied to predict heart disease. Literature survey is then covered in paper providing methods adopted by different researchers. In research methodology author had given detailed knowledge of multi linear regression its mathematical model, equations and how it can be applied to predict heart disease and author concluded paper by giving some experimental results in tabular and chart form, conclusion and some references that author had referred to.

4. Experimental work:

- Objective:
  Prediction of Heart Disease using Multiple Linear Regression Model
- Dataset
  - Heart disease dataset by UCI to train machine. Dataset contain values like Sex, Chest Pain Type, Fasting Blood Sugar, Restecg, Exang, Slope, number of major vessels colored by floursopy, Thal, Trest Blood Pressure, Serum Cholesterol, Thalach – maximum heart rate achieved, ST depression and Age.

- Techniques:
  - Linear regression,
  - Logistic regression and
  - Artificial neural network.

- Results:
  - Analysis shows that whether the patient has heart disease or not.

5. Conclusion:

Today diagnosing patients correctly and administering effective treatments have become quite a challenge. The diagnosis of diseases is a critical and complicated job in medicine. The cost to treat a patient with a heart problem is quite high and not affordable by every patient. The recognition of heart disease from diverse signs is a mysterious problem that is encountered with number of false assumptions and is frequently accompanied by impulsive effects. So, there is a need to present an efficient approach for extracting significant patterns from the heart disease data warehouses for the efficient prediction of heart attack. Hence, Multiple Linear Regression Analysis is performed on trained data to build a model on which test data is applied. From the experimental results it is proved that Multiple Linear Regression is appropriate for predicting heart disease chance.

6. References:

- Prediction of Heart Disease using Multiple Linear Regression Model by K.Polaraju and D.Durga Prasad.
- Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-48.
- Anooj, P. K. (2012). Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. Journal of King Saud University-Computer and Information Sciences, 24(1), 27-40.
- Parisa Naraei, Abdolreza Abhari and Alireza Sadeghian," Application of Multilayer Perceptron Neural Networks and Support Vector Machines in Classification of Healthcare Data", IEEE, 2016.
- DeepaliChandna "Diagnosis of Heart Disease Using Data Mining Algorithm", IEEE Conf. on International Journal of Computer Science and Information Technologies, 2015, pp 1678-1680.
- Asghar, S. "Automated Data Mining Techniques: A Critical Literature Review" 978-0-7695-3595-1, 75 – 79, IEEE, 2009.

- M.Akhiljabbara“Heart Disease Prediction System using Associative Classification and Genetic Algorithm”IEEE, 2012.
- Niti Guru, Anil Dahiya, Navin Rajpal, “Decision Support System for Heart Disease Diagnosis Using Neural Network”, Delhi Business Review, Vol.8, No.1, 2007
- K. Srinivas, B.Kavitha Rani and Dr. A. Govrdhan “Application of Data Mining Techniques in Healthcare and Predication of Heart Attacks”, International Journal on Computer Science and Engineering, Vol. 02, No. 02, pp.250- 255,2011.
- N. Deepika and K... Chandrashekar, “Association rule for classification of Heart Attack Patients”, International Journal of Advanced Engineering Science and Technologies, Vol.11, No.2, pp253-257, 2011.
- D. Shanthi, G.Sahoo and Dr. N. Saravanan, “Designing an Artificial Neural Network Model for the Prediction of Thrombo-embolic Stroke”, International Journal of Biometric and Bioinformatics, Vol. 3, No.1,pp250-255,2008.