**International Journal Of Engineering Research And Development**
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 14, Issue 6 (June Ver. I 2018), PP.01-06

# Heart Disease Prediction By Machine Learning

Dhanushree.Y, Indrani.N, Santosh.N, Vidya Vani.R, Dr.Rajeshwari.J
*Department of Information Science & Engineering, Dayananda sagar College of Engineering, Bangalore.*
*Corresponding Author: Dhanushree.Y*

**ABSTRACT:-**The growth of big data in the healthcare field enables easy and early prediction of the disease and ensures the patients care. As there are different characteristics for different diseases in the different regions, the prediction of the disease outbreaks will not be so accurate so machine learning algorithms are streamlined for predicting the chronic diseases effectively and to solve the difficulty of incomplete data. The modified prediction models are experimented over the real life hospital data for the disease prediction.
**KEYWORDS:-** Diseases, Hospitals, Prediction algorithms, Big Data, Data models.

---

Date of Submission: 25-05-2018 Date of acceptance: 09-06-2018

---

## I INTRODUCTION

Data mining plays an important role in various applications such as business organizations, e-commerce, health care industry, scientific and engineering. In the health care industry, the data mining is mainly used for Disease Prediction. The objective is to predict the diagnosis of heart disease with reduced number of attributes. Here fourteen attributes involved in predicting heart disease. But fourteen attributes are reduced to five attributes by Naïve Bayes classifiers.Heart disease is a general name for a variety of diseases, conditions and disorders that affect the heart and the blood vessels. Symptoms of heart disease vary depending on the specific type of heart disease. Congenital heart disease refers to a problem with the heart's structure and function due to abnormal heart development before birth. Congestive heart failure is when the heart does not pump adequate blood to the other organs in the body. Coronary heart disease or in its medical term Ischemic heart disease is the most frequent type of heart problem. Coronary heart disease is a term that refers to damage to the heart that happens because its blood supply is decreased, it leads to the fatty deposits build up on the linings of the blood vessels that provide the heart muscles with blood, resulting in them narrowing The paper identifies the risk factors for the different types of heart diseases.

Most hospitals today employ some sort of hospital information systems to manage their healthcare or patient data .These systems typically generate huge amounts of data which take the form of numbers, text, charts and images. Unfortunately, these data are rarely used to support clinical decision making. There is a wealth of hidden information in these data that is largely untapped.The main aim is the effective prediction of chronic diseases such as heart diseases at the earlier stage and reduces the cost of medical expenses. This can be achieved by streamlining Naive Bayes machine learning algorithm with the reduced set of attributes and efficiency will be increased by the precise prediction at the earlier stage.

## II LITERATURE SURVEY

In past years many research and studies have been carried out to accelerates the alue and innovation to give new ways of knowledge to healthcare stakeholders. This enables various changes in quality of healthcare and also healthcare expenses will be minimized [1]. Mining electronic health records towards better research applications and clinical care was established for undiscovered classification of patient principles and for detecting the unknown diseases. Data accessibility accuracy and thoroughness of massive data is easy to explore and thus provides better clinical care applications [2]. Risk factors and risk assessment tools were proposed for falls in hospital in-patients in a systematic review to identify individual clinical risk factors and the tools for clinical risk assessment. The falls can be prevented to maximum extent by the qualitative fall prevention facilities[3].Data mining for censored time-to-event data which is a Bayesian network model was proposed for predicting cardiovascular risk from electronic health record data predict whether the cardiovascular event will be occurred within five years by using Bayesian network model. By using the predictive models for censored data, the performance criteria will be increased [4]. Big data in health care using analytics was proposed to identify and manage high-risk and high-cost patients to recognize how connection of humans is leading to increase data. It is the technology that will connect the human and help them acquiring innovative knowledge [5].

## III METHODOLOGY

The system is developed to register the details of the patient to full his details which undergo certain classification which mainly consist of the following methods:-

3.1 Attribute to Weight Conversion
3.2 Average Weight Calculation
3.3 Excel Data Uploading
3.4 MVC architecture

**3.1 Attribute to Weight Conversion:**
Each attribute is assigned a weight ranging from 0-1 to reflect their importance in prediction of disease.

**3.2 Average Weight Calculation:**
A simple mathematical model that accurately predicts individual weight change offers opportunities to understand how individuals lose and gain weight and can be used to foster patient adherence to diets in clinical settings.

**3.3 Excel Data Uploading:**
The large data available from medical diagnosis is analyzed and useful information known as knowledge is extracted by using data mining tools.

**3.4 MVC architecture:**
It is an architectural pattern commonly used for developing user interfaces that divides an application into 3 interconnected parts.

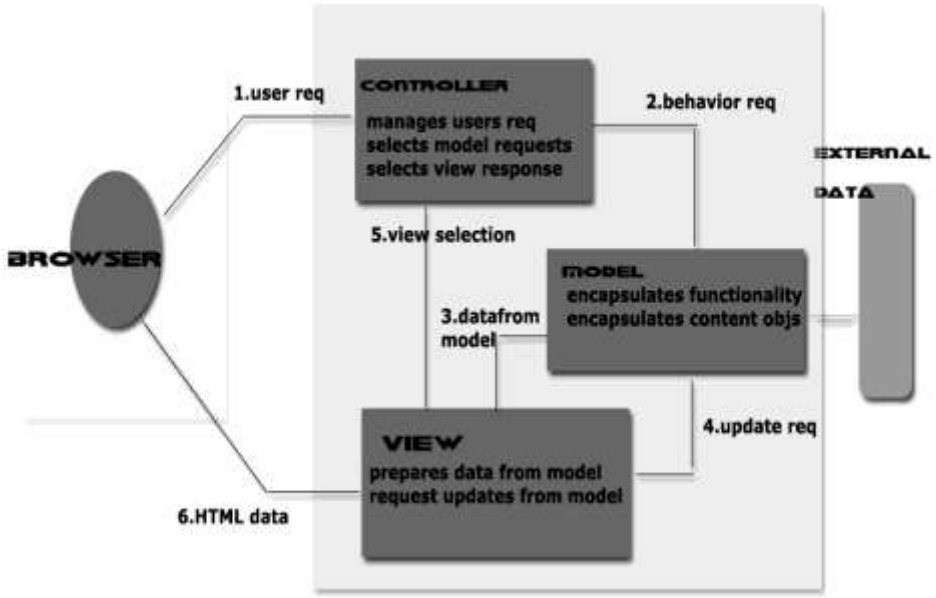



**Figure:-MVC Architecture.**

## IV SYSYTEM DESIGN AND ARCHITECTURE

The system architecture consists of two modules admin login and the user login. Admin will upload data from the excel sheet. This data will undergo the weightage conversion and the records are stored in data base and this data is clustered for precise prediction. This clustered data is then retrieved from the data base and undergoes Naive Bayesion classification and then will display the result. This result will sent to the user and the user will convert the input details into weightage conversion by Naive Bayesion classification and the result will be displayed.

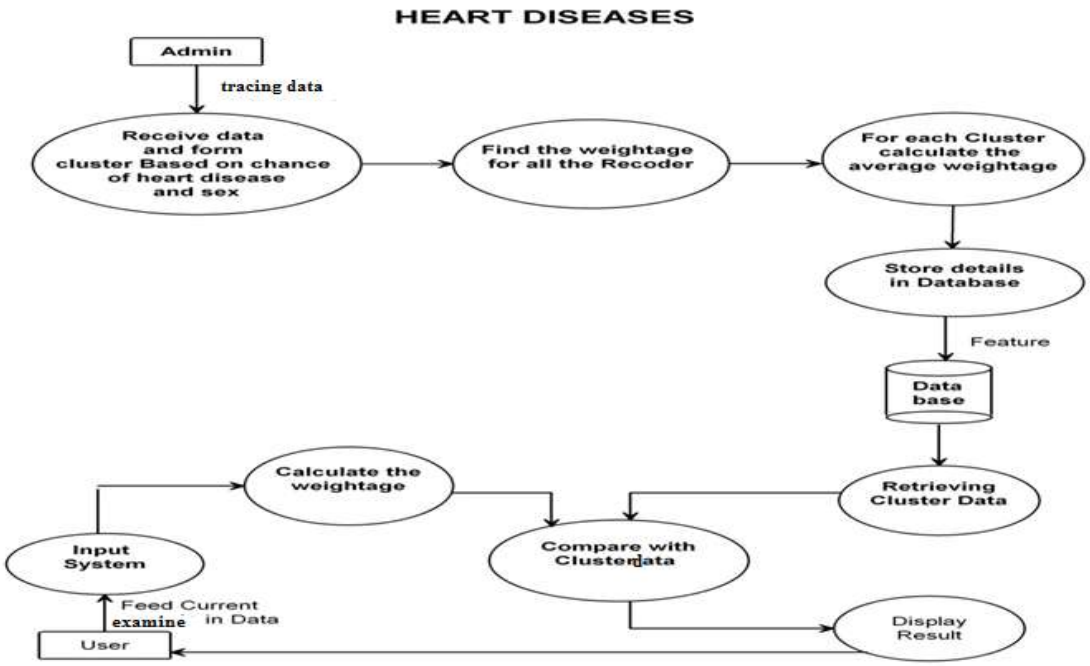



**Figure: System architecture using Naïve Bayes classifiers.**

**4.1 Design Goals**:-The following are the design goals that are applicable regardless of application domain, size, or complexity.
4.1.1Simplicity
4.1.2Consistency
4.1.3Identity
4.1.4Visual appeal
4.1.5Compatibility

Design leads to a model that contains the appropriate mix of aesthetics, content, and technology. The mix will vary depending upon the nature, and as a consequence the design activities that are emphasized will also vary.

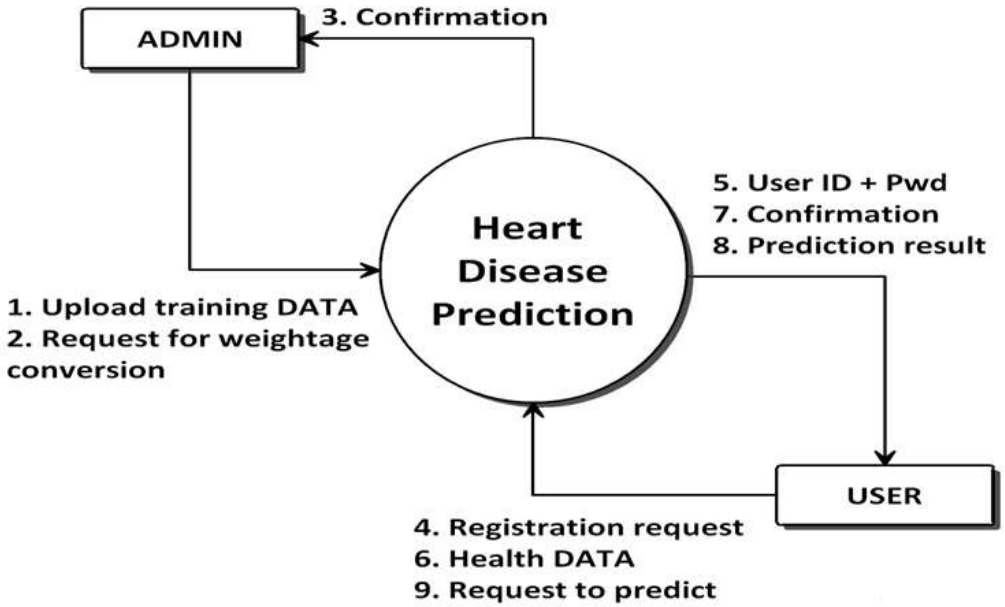
**4.2 Advantages:-**
1. Early prediction of heart diseases can be done.
2. The cost of medication will be minimized.
3. Accuracy rate will be high.

## V IMPLEMENTATION

**5.1 Patient Details**
1. View patients details
2. Add patient details
3. Update patient details
4. Delete patients details

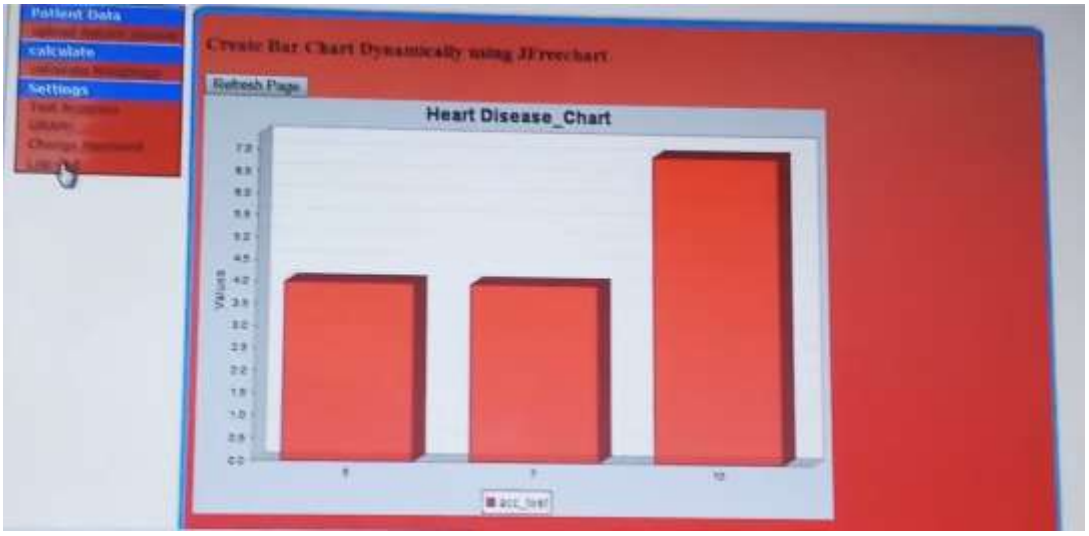
**5.2 Patient Records**
1. Add patient records
2. View patient records

**5.3 Settings**
1. Change password
2. Logout




## VI RESULTS


1. Uploading patient data




2. Calculating accuracy by graph


3. Display of final report

## VII CONCLUSION

The main aim is to experiment the training data over the hospital data and effectively predict the disease at earlier stage by using Naive Bayes machine learning algorithm using the concept of reduced set of attributes and weightage conversion. To the best of our knowledge, none of the existing work focused on both data types in the area of medical big data analytics. Compared to several typical prediction algorithms, our proposed algorithm reaches high accuracy.

## REFERENCES

[1]. P. Groves, B. Kayyali, D. Knott, and S. van Kuiken, The`Big Data 'Revolution in Healthcare: Accelerating Value and Innovation. USA:Center for US Health System Reform Business Technology Office, 2013.

[2]. P. B. Jensen, L. J. Jensen, and S. Brunak, ``Mining electronic healthRecords: Towards better research applications and clinical care," Nature Rev. Genet., vol. 13, no. 6, pp. 395_405, 2012.

[3]. M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, ``Wearable 2.0:Enable human-cloud integration in next generation healthcare system,"*IEEE* Commun., vol. 55, no. 1, pp. 54_61, Jan. 2017.

[4]. D. Oliver, F. Daly, F. C. Martin, and M. E. McMurdo, ``Risk factors and risk assessment tools for falls in hospital in-patients: A systematic review, 'Age Ageing, vol. 33, no. 2, pp. 122_130, 2004.

[5]. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, ``Big data in health care: Using analytics to identify and manage high-risk and high-cost patients," Health Affairs, vol. 33, no. 7, pp. 1123_1131,2014.

[6]. S. Bandyopadhyay *et al.*, ``Data mining for censored time-to-event data:A Bayesian network model for predicting cardiovascular risk from electronic health record data," Data Mining Knowl. Discovery, vol. 29, no. 4,pp. 1033_1069, 2015.
[7]. B. Qian, X. Wang, N. Cao, H. Li, and Y.-G. Jiang, ``A relative similarity based method for interactive patient risk prediction," Data Mining Knowl.Discovery, vol. 29, no. 4, pp. 1070_1093, 2015.
[8]. N. Nori, H. Kashima, K. Yamashita, H. Ikai, and Y. Imanaka, ``Simultaneous modeling of multiple diseases for mortality prediction in acute.

Dhanushree.Y. "Heart Disease Prediction By Machine Learning." International Journal Of Engineering Research And Development , vol. 14, no. 06, 2018, pp. 01–06