

# E-Smart Health Prediction System using Datamining Tools

Professor Dr. Bindu Garg, Aamir Hafiez, Shalini, Mayank Kumar

Dept. of Computer Engineering,

Bharati Vidyapeeth Deemed to be University College of Engineering, Pune, Maharashtra.

**Abstract** – This project focuses on medicinal field, educational field, research field and various aspects of human’s physical and mental well-being. Due to the advancement of technology all over the world, availability of computers and digital data is in abundance. This data can be used for various purposes to deal with the modern medical problems because such data is not just collected from a single individual rather by a huge data set of individuals. But there is no need to deal with such a huge amount of data at a single time rather only a part of data is actually required for a certain use. Hence the main objective is to use data mining tools to predict an accurate result that can help to improve various aspects of Clinical Predictions which in turn will help society and doctors to get an automated tool to easily predict the disease. The major objective of this paper is to evaluate data mining tools and techniques in medical field to develop an accurate disease predicting application. Data mining is a totally new and advanced field of Computer Science and Engineering which focuses on large data to extract only useful data sets. It is a process of discovering patterns in large data sets involving techniques of Database systems, artificial intelligence, machine learning and statistics. Its overall goal is to extract information with intelligent methods. Thus with the help of data mining an intelligent software can be developed which will be helpful for all the society members.

**Keywords** – Smart Health Prediction System, Data Mining, Clinical Predictions.

## I. INTRODUCTION

Health is Wealth. A healthy person is always happy in his life. People with good health are one who always nurture the world around them. People who are unfit can never enjoy the sweet fruit known as life. A healthy body resides a healthy mind. Due to the advancement in technology the health industry has grown a lot from past years. It is shown that at least five terabytes of data is generated from a health care hospital every year. Hence it is not a miracle to know that this much amount of data can be used to predict another new patient who has the same following symptoms that preceding patients had last year or maybe other past years. Thus data mining techniques can be very useful to extract this information based on the person’s new symptoms.

This novel computer engineering technique hence will be able to accurately predict the correct disease and hence will be able to diagnose, prevent, and cure major as well as minor illnesses. The outcome of the data mining application here is to provide benefits to health care organizations for grouping the patients having similar type of disease or health issues so that health care organization provides them effective treatment. It can also be used for management of patient’s duration of stay, or provide cost management facility, and making plans for effective health care management. WHO has defined E-health as, the use of information and communication

technologies (ICT) for health to, for example, treat patients, pursue research, educate students, track disease and monitor public health. Recently a new virus has emerged around the globe known as COVID – 19. This proposed system will help to extract information from its large data sets, like generated from previous coronaviruses such as SARS and MERS and calculate similar patient symptoms and give a proposed treatment for this novel coronavirus. The proposed system is mainly used by certain people to provide integrity and confidentiality of data. Data mining has variety of scopes in major fields such as:

- Administration of health services;
- Clinical care;
- Medical research; and
- Training.

## II.LITERATURE SURVEY

In this paper we propose to identify mining algorithms that are efficient to give us the required output. We can create an application which will take into consideration of various factors that affect the health of a particular individual by using predictive analytics and data mining tools and techniques.

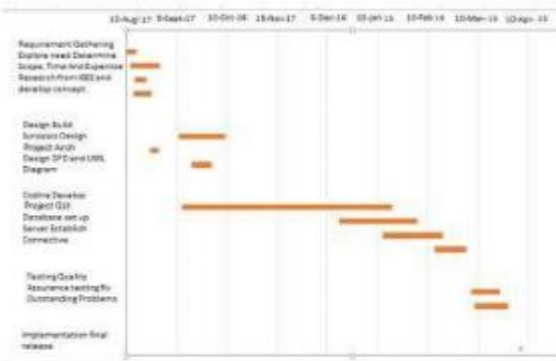
Naïve Byes Algorithm is used to compare frequent data items with existing algorithms.

Data mining techniques are used on the medical data which has abundant scope on improving health solutions.

When used correctly, electronic health data and records will be very useful and prove miraculous in medical treatment sector. The data that is generated by the health care sector is complex and consists details about patients, diseases, diagnostics, resources, medications, treatments etc.

### III. PROPOSED SYSTEM

It often happens that any of us require treatment or diagnosis and hence require the help of doctor immediately. There may be certain instance where the doctor may not be available or the situation can be that we ourselves are not able to access any doctor nearby and thus can rely on the proposed system. This system will be live and put online. Hence entire health monitoring system will be online and ready to use by any individual around the globe. The main feature of the system is similar to a doctor, where the doctor asks the symptoms of the patient and predicts the disease that the patient is/maybe suffering form and thus prescribe him with the required treatment. The system predicts the disease based on the symptoms that the user will enter. The system analyzes the patient's symptoms and based on that searches in its huge data set for similar trends of data and predicts the best outcome for the same. In this system we use the data mining techniques to extract the most accurate disease or illness that could be related with patient's symptoms.



If the system is unable to process the symptoms to provide a solution, then it will inform the user the probable disease he has.



Fig .2.Screenshot.

If the symptoms entered by the user does not match with any of the diseases present in its database then the system will propose the most probable disease that the patient is suffering based on his symptoms. This system tends to replace the need of doctor by proposing smart solutions for the diagnosis of disease or illness that the user may be suffering from. The actual feature of the system is to give an instant diagnosis of the disease based on the symptoms entered by the user. In the system data mining algorithms are used to cross check the database and from that frequent item sets are mined out.

### IV. ALGORITHM USED

#### 1. Naïve Byes –

Naïve Byes is one of the most popular and simplest machine learning classification algorithm. It is based on Bayes theorem for calculating probabilities and conditional probabilities. It is a very extremely fast algorithm relative to other classification algorithms and works on the principle of Bayes theorem of probability to predict the class of unknown data sets. Naïve Bayes algorithm rather being a single algorithm is a family of algorithms based on common principle. Naïve Bayes classifier assumes that the presence of a particular feature in a class is not related to presence of any other feature. For example, in a basket of fruits, an apple is red, round and approximately of three inches in diameter. All these feature may depend on each other or upon existence of other features, but they independently contribute to the probability that the fruit actually is an Apple.

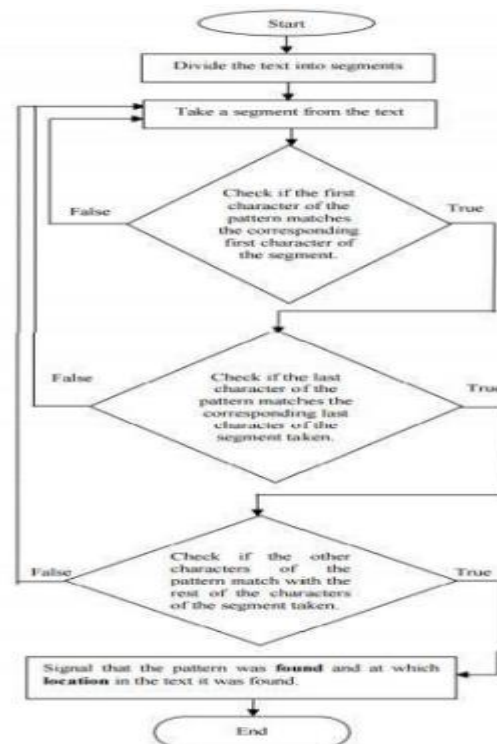


Fig .3.Naïve Bayes flowchart.

Above is the flowchart shown how Naïve Bayes algorithm workflow actually is.

## 2. Naïve Bayes Algorithm Steps –

- Step 1: Calculate prior probability for given class.
- Step 2: Calculate conditional probability with each attribute for each class.
- Step 3: Multiply same class conditional probability.
- Step 4: Multiply prior probability with step 3.
- Step 5: See which class has higher probability, higher probability class belongs to given input set step

## V. RESOURCE USED

### 1. Hardware:

- Ram : 2 GB.
- Rom : 100GB.
- Processor : Minimum 2.0 GHz.

### 2. Software:

- OS : Windows 10
- Language : Java, JDK 1.7
- Database : MySQL
- IDE : Eclipse
- Server : Apache Tomcat 2.0

## VI. CONCLUSION AND FUTURE WORK

In the proposed system hidden knowledge will be extracted in order to predict future outcomes accurately. This system will be very helpful where the availability of doctor may be nil or minimum. The system proposed will be easily able to predict the accurate results since it will be based on actual historical data, which will help patients to be diagnosed with the disease or illness immediately without the consultation of a doctor. The future work that can be proposed for the system will be to provide an online doctor which will provide a real time help for the patients and help with their concerns. Another enhancement that can be made for the proposed system can be that an android application can be made so that users can easily access the application remotely.

## VII. ACKNOWLEDGEMENT

We express our deep gratitude to our project guide (Dr. Bindu Garg) for providing timely assistance wherever and whenever required. She has been a lighthouse to journey of making of this project.

We extend our gratitude to our Head of Department, Dr. Devendra Thakore who gave us this opportunity to work on this project and gave us every possible time and resource for its completion.

We would also like to thank all the faculty members of our department for their valuable insight and tips in designing the system.

## REFERENCES

- [1]. Thareja, Reema, “Data Warehousing”, 2009.
- [2]. Dean, Jared, “Big Data, Data Mining and Machine Learning”, 2014.
- [3]. M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.
- [4]. International Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2017 IJSRCSEIT | Volume 2 | Issue 2 | ISSN : 2456-3307. M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang,
- [5]. C. Youn, “Wearable 2.0: Enable Human- Cloud Integration in Next Generation Healthcare System,” *IEEE* Jan. 2017.