Health Prediction by Data Mining

Rohan Lohia[1] | Vibhor Sharma[2]

[1] B.Tech Scholar , Department of IT, Maharaja Agrasen Institute Of Technology, Delhi, India
[2] Assistant Professor, Department of IT, Maharaja Agrasen Institute Of Technology, Delhi, India

ABSTRACT

The paper presents an overview of the Clinical Predictions and Medical Predictions with data mining and its techniques. In health care areas, due to regulations and due to availability of computers, such large amount of data cannot be processed by humans to schedules and diagnosis in short time of duration. It is a new technology which is of high interest in computer world. The computer world make an data in different databases to transfer into new researches and results. The database management extract a new patterns from large datasets. The different parameters included in data mining are: clustering, forecasting, path analysis and predictive analysis.

KEYWORDS: *Health Prediction, Data Mining, Clinical Predictions, Clustering, Forecasting, Predictive Analysis, Semi-Automatic Means.*

I. INTRODUCTION

The health care industry has been growing from past few years. So in order to overcome such problem we have designed a user friendly website which helps to identify their problem from their residence at any time. This websites provide an option for booking an appointement with doctor to discuss health related problem and get diagnosed properly.

This data mining and its technique is used to analyze a large amount of data. The Clinical Predictions and medical predictions has a warehousing in combination with data mining which can help administrative, clinical, research and education aspects of health institutions. The data mining has variety of scopes and some of which are listed below:

*Administrative of health service
*Health care
*Medical research
*Training

There are various challenges of data mining which can be discussed in health prediction. There are currently lot of health institutions that has been developed such as hospitals and medical centers which are vital to maintain and improve the health community around us. For every illness and diseases that people may face today or in future, it is because of these unhealthy environment or food and all the doctor who worked at hospitals or clinical center that have made our lives physically better and healthy. Although hospitals now are well-equipped but there are still issues that persists the staffs to make poor clinical decisions that affects a patient's health such as lack of health

information, lack of qualified doctor and communications between doctors and patients.

With all of these issues mentioned, the solutions are needed to be made and that is, the patients can check their diseaase by online booking appointment for their several disease which can checked by the doctor. Smart Health Prediction should be implemented which could potentially eliminate these concerning issues.

LITERATURE SURVEY

In this era, Health care institution are very essential for every single people in this world as a proper health care. This institution purpose is to improve the current health care community that have shared and created. A health care institution such as hospitals or medical centers would have consists of numerous of doctors that were qualified and have a specialize on their fields which can treating patients of their current illness and to restore them to proper health.

Nowadays, new technologies have been created and developed to improve people's life and routine especially for health care. Few years ago, doctors were expected to use their own experience to handle every medical situation that patients are facing every day. Doctors and Nurses were now guided by smart health prediction of storing information and tested that may be used for research and diagnosis. Although their current approaches may have saved people's live back that have endangered the human life by prone to errors and wrong doings. It's no doubt that heavy decision or heavy burden for everyone especially the medical staffs to understand that a number of decisions could heavily affect other people's lives and health, it is also why such institutions has proves to be a vital on guiding their medical staffs to make a proper clinical decision to cure and save the human health.

A smart health prediction is defined as a health care system that is about to assist health professionals in their decision-making process regarding to medical situations. It will provide the guidance and information needed for doctors and patients on medical illness and it will eliminate the difficulties that the patients needed a professional doctor to cure their illness without any wrong decision and doctors needed to particularly in their clinical decision-making process. This system would require a lot of medical information that are valuable to be used on predicting a patient's health status, these information will be analyzed by using data mining techniques in order to find a new pieces of information from unstructured data. By using data mining tools, it will not only be able to produce reliable results with less time consumption and complexity but also with smart decision and useful information. As we all know this ongoing pandemic (COVID-19), all people's are afraid of this corona virus . This virus spread by touching or close-gathering to the society and their main symptoms are cold & cough, high fever and throat pain. People's are so conscious that they may cure by itself and the smart health prediction with this disease to cure and treating people's by just type the symptoms and the disease then they may get a medicine name or wanted to appoint a doctor for their treatment.

CONTENT

This section explains the data mining process and its application in health prediction. It also analysis the related application of data mining techniques in health prediction.

Data Mining

The health prediction will depend on its implementation of data mining which is referred as data mining knowledge and information from a large amount of data sets. The institutions are one of the vast amount of information that can utilized helpfully by data mining of many fields in society. Data mining can improve the medical center or health care institution with eliminating current health by easily providing to complex medical cases to solve and eliminate any time consumptions created and making a clinical decision.

Data mining is described as the process of searching certain patterns that comes in a database and utilize that information to predictive models. Its process involve examining and selecting certain data from a data storage to uncover new and unknown patterns.
Data mining is also known process of involving in Knowledge Discovery in database which is the extraction of large amount of data from databases. It uses to examine hidden patterns and relationship that can be found a large amount of data from decision-making purposes. Here, Knowledge Discovery in Database is of 7 sequential steps as follows:-
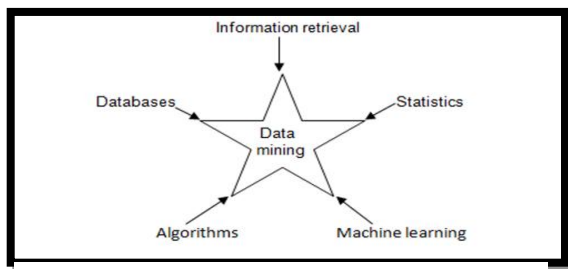
Data Cleaning
It is represented as the first step of KDD which requires to eliminate any data collected that is missing or irrelevant.

Data Integration
It is the second step which takes the data that has been filtered from the previous step to combined into a meaningful and useful data.

Data Selection
It is defined as the process of which data that is relevant for the analysis is selected and recover from the collection of data.

Data Transformation
It involves the data being converted into forms that is required for performing different mining technique such as smoothing or normalization.

Data Mining
It consists of examining the data for any pattern or rules that are useful for extracted and obtained.

Pattern Evaluation
It is defined as identifying the data for any patterns that represents the knowledge based on specified measure.

IMPLEMENT
The implementation of a health prediction system will allow doctor and medical staffs to relieve their efforts on their medical or clinical decision-making process by simply add the user's health data and symptoms that they are suffering. This implemented with data mining technique that may deduce the disease by correlating the information given by the patients with their health information for the doctor and medical staffs provides and then they stored the data in a database. The entire process would efficiently reduce the time ruin and challenges efforts that doctors put themselves into for making a medical decision.

This project is delibrate to deliver a user-friendly for the patients and doctors to use diagnosing illness and to provide suitable guidance on their health isse that they are facing.

This system needs to include the various functionality:

Patient Registration
Patients would require registration by themselve for the first time login with their username and password to use the websites.

Patient Login
Patient would require to login with username and password.

Viewing Patient Details
Doctors and patients may view details of one another by themselve.

Disease Prediction
The user is trying to describe by several questions that they are facing the illness/disease using data mining to the most accurate symptoms.

Search Doctors/Patients
Doctors and Patients may search according to their speciality, disease that they had contracted and references.

Adding Disease and Symptoms
Administrators may add new disease and symptoms into the website/system for doctors and patients.

Doctor Login
Doctors are required to login with their username and password to use the access.

Doctor Registration
Administrators can add and register a new doctor into the system and add their ussername and password to access the login.

Admin Login
Only doctor can access the admin login themselve with their username and password.

View Diseases
Administrator can view various diseases details which are stored in the database of the system.

Providing Feedbacks

Doctors and patients may provide feedbacks that may serve as an additional information to be viewed.

Sharing Information

Doctors can share the information of a diseases or patients to another doctor for verification.

CONCLUSION

The data mining can play an vital role in disease to health prediction. The combination of several data mining techniques may be a better approach in designing health prediction website. This study does not enclose thecomplete analysis of all the existing data mining technique and a real-time healthcare database. The future research may be directed towards the selection of the best suitable data mining technique through all the existing analysis.

REFERENCES

[1]" Survey on data mining techniques for diagnosis of diseases in medical domain",R siddarth, I parvathi international journal of computer science and information technologies, 2014..

[2] S. Chowdhury, "A Review Paper on Health Data Integration with Secured Record Linkage in National Health Data Warehouse and Privacy & Compatible Solution for its Privacy and Security Problems in Bangladesh and other developing countries," International Islamic University Chittagong (IIUC), 2017.[https://works.bepress.com/sayem-binsarwarchowdhury/2/download/.]

[3] V. Kirubha and S. Priya, "Survey on data mining algorithms in disease prediction,"*International Journal of Computer Trends and Technology (IJCTT).* [https://pdfs.semanticscholar.org/0d05/05b5b057333888de85043ae9b41259f07d85.pdf.]

[4] K. Lakshmi, D. Ahmed and G. Kumar, "A smart clinical decision support system to predict diabetes disease using classification techniques," *IJSRSET.*

[5] B. Thakkar, M. Hasan and M. Desai, "Health care decision support system for swine flu prediction using naïve Bayes classifier,"*IEEE International Conference on Advances in Recent Technologies in Communication and Computing.* [https://sci-hub.tw/https://ieeexplore.ieee.org/abstract/document/5656881.]

[6]Mohammed, Zaki. J, Meira, Wagner,"Data Mining and Analysis: Fundamental Concepts and Algorithm", 2014.

[7] D. KS and A. Kamath, "Survey on Techniques of Data Mining and its Applications,"*International Journal of Emerging Research in Management & Technology,* [https://pdfs.semanticscholar.org/b738/3df4705133a132f58104b514b80555fe78cb.pdf]