DYNAMIC SYMPTOMS BASED DISEASE DETECTION WITH DRUG PREDICTION USING BIG DATA AND MACHINE LEARNING

ISANAKA SAI VENKATESH REDDY, S, MAGESH KUMAR

UG Scholar, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

Associate Professor, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

venky11.05.1999@gmail.com, mageshkumars.sse@saveetha.com

ABSTRACT

Our goal is to provide a tool to assist professionals and consumers in finding and choosing drugs. To achieve this goal, we develop an approach that allows a user to query for drugs that satisfy a set of conditions based on drug properties, such as drug indications, side effects, and drug interactions, and also takes into account patient profiles. We also analyze the disease and best drug advised to that specific patient through Big Data analysis we are implementing an application to know about disease based on our symptoms. Based on the symptoms system will predict the type of disease and suggest the specialist doctor based on the rank method. After that best drug will suggested by the system. Order based Data mining plays significant job in different social insurance administrations. In social insurance province, significant as well as testing task are to analyze wellbeing circumstance as well as appropriate therapy of ailment at the advance arrange. Some of different ailments that can be analyzed advance also, can be attend to at the beginning period. While infections. The customary methods for diagnosing thyroid infections relies upon clinical assessment as well as number of tox screen. Important assignment in order to recognize malady conclusion at advance phases accompanied by high precision. Data mining strategies take part in an significant job in social insurance area for creating decision, illness finding as well as giving best therapy to sufferer effortlessly. Sickness categorization is an significant job. Motivation behind investigation is to predict the illness utilizing distinctive arrangement procedures.

Keywords: classification model, Data mining tools, Decision tree, Illness diseases, drugs, predictive analysis

I. INTRODUCTION

Big data is an all-encompassing term for any collection of data sets so large and complex that it becomes difficult to process using traditional data processing applications. The challenges include analysis, capture, duration, search, sharing, storage, transfer, visualization, and privacy violations. The trend to larger data sets is due to the additional information derivable from analysis of a single large set of related data, as compared to separate smaller sets with the same total amount of data, allowing correlations to be found to ”spot business trends, prevent diseases, combat crime and so on. So we can implement big data in our project because every employ has instructed information so we can make analysis on this data. Big data is a term utilized to refer to the increase in the volume of data that are difficult to store, process, and analyze through traditional database technologies. The nature of big data is indistinct and involves considerable processes to identify and translate the data into new insights. The term “big data” is relatively new in IT and business. However, several researchers and practitioners have utilized the term in previous literature. For instance, referred to big data as a large volume of scientific data for visualization. Several definitions of big data currently exist.” Meanwhile and defined big data as characterized by three V s: volume, variety, and velocity.
II. OBJECTIVE OF THE PROJECT:
The objective of the project is to identify the disease based on the patient symptoms and suggest the best drug by English and ayurvedic medicine. To suggest doctor based on ranking.

III. LITERATURE SURVEY:
Diagnosing nosocomial infections in critically ill patients admitted to intensive care units is a challenge because signs and symptoms are usually non-specific for a particular infection. In addition, the choice of treatment, or the decision not to treat, can be difficult. Models and computer-based decision-support systems have been developed to assist ICU physicians in the management of infectious diseases. We discuss the historical development, possibilities, and limitations of various computer-based decision-support models for infectious diseases, with special emphasis on Bayesian approaches. Although Bayesian decision-support systems are potentially useful for medical decision making in infectious disease management, clinical experience with them is limited and prospective evaluation is needed to determine whether their use can improve the quality of patient care. This violates the privacy-preserving requirements for some legal or commercial reasons. The problem is that the classifier learned by the inherently violates the privacy. This privacy violation problem will restrict the applicability of the SVM. To the best of our knowledge, there has not been work extending the notion of privacy preservation to tackle this inherent privacy violation problem of the classifier. In this paper, we exploit this privacy violation problem, and propose an approach to post process the SVM classifier to transform it to a privacy-preserving classifier which does not disclose the private content of support vectors. The post processed classifier without exposing the private content of training data is called Privacy-Preserving SVM Classifier. The PPSVC is designed for the commonly used Gaussian kernel function. It precisely approximates the decision function of the SVM classifier without exposing the sensitive attribute values possessed by support vectors. By applying the PPSVC, the classifier is able to be publicly released while preserving privacy. We prove that the PPSVC is robust against adversarial attacks. The experiments on real data sets show that the classification accuracy of the PPSVC is comparable to the original classifier. Electronic healthcare is becoming a vital part of our living environment and exhibits advantages over paper-based legacy systems. Privacy is the foremost concern of patients and the biggest impediment to e-healthcare deployment. In addressing privacy issues, conflicts from the functional requirements must be taken into account. One such requirement is efficient and effective response to medical emergencies. In this article, we provide detailed discussions on the privacy and security issues in e-healthcare systems and viable techniques for these issues. Furthermore, we demonstrate the design challenge in the fulfillment of conflicting goals through an exemplary scenario, where the wireless body sensor network is leveraged, and a sound solution is proposed to overcome the conflict. E-healthcare systems have been increasingly facilitating health condition monitoring, disease modeling and early intervention, and evidence-based medical treatment by medical text mining and image feature extraction.

Problem statement:
In the Existing system, more number of people were getting serious and even died because of not aware of disease. Even more number of applications are available to know about disease but they need login and patients have to give symptoms manually. Regularities in the prescriptions are for both clinical practice and the novel prescription development.

Proposed System:
We implement a web application to know about their disease themselves using Big data. People were easily gives their symptoms by selecting symptoms form drop down box. we develop an approach that allows a user to query for drugs that satisfy a set of conditions based on drug properties, such as drug indications, side effects, and drug interactions, and also takes into account patient profiles. We also analyze the disease and best drug advised to that specific patient through Big Data analysis we are implementing an application to know about disease based on our symptoms. Based on the symptoms system will predict the type of disease and suggest the specialist doctor based on the rank method. After that best drug will suggested by the system. Based on disease we suggest the medicine in allopathic and ayurvedhic.
IV. ARCHITECTURE DIAGRAM:

![Diagram showing the flow of data from patient to system]

V. DATA FLOW DIAGRAM:

![Diagram showing the interaction between user and system]

Results and Discussion:

VI. MODULES:

1. SERVER DEPLOYMENT
   - Data Service Provider will contain the large amount of data in their Data Storage. Also the drugs and disease information will be stored in the Database. Also the Data Server will redirect the User requested job to the Resource Assigning Module to process the User requested Job.

2. CONSTRUCTION OF DISEASE TRAINING SET
   - In this module we can design and implementation of train the disease to system. Server will store a set of trained dataset and its relevant diagnosis pattern. Multiple number diseases and symptoms will be collected and stored in a training set. Different type diseases and symptoms will be stored in a dataset.
3. BIG DATA BASED ANALYSIS

In this module we implement big data, in this big data we will have lot or vast amount of data that may wanted or unwanted information in simple the information in the big data are unstructured. So in this module the patient is going allow permission to access the server by the big data analyst. The big data analyst get the all the disease and drugs information which mention above and extract the information by the technique of map reducing formation to get useful information which is useful for patient.

4. PRE STORED DATA COMPARISON

In this module doctor will import all the details about the medicine i.e. what are the symptoms, dosages and drug .And hw will store more about of the medicine so that we can make some use of it for example we can give awareness to the society .we store the all the data in the clustering format so that data can spitted and stored in the different clusters. So that it will easily to classify the data for the research.

5. PREDICTIVE DISEASE ANALYSIS

In this module we implement predictive disease analysis system in which the data will be analysis so that we can predictive the disease based on the symptoms .This module interact with server to analysis, the analysiation is done by the researchers. So they get the data from the server to make analysis to find the disease based on the symptoms

6. BEST DOCTOR RECOMMENDATION

In this module we suggest the best doctor based on ranking. These ranking system is based on patient review. Patients will give review for the treatment and based on review doctors will come upon top most place.

7. DRUG AND TRAINING SET CONSTRUCTION

In this module we will train the drugs for every disease and also train the side effects of the drugs. User will be giving their Symptoms & Diagnostic reports to the system for the diagnosis of the disease. If any side effects came by using the prescribed drug user can give the side effects on the same website and those side effects will be match the drugs and suggest the alternate drug for that disease. In this we will suggest the medicine Based upon patients willing like whether they need ayurvedic or English medicine.

VII. IMPLEMENTATION SCREENSHOTS
CONCLUSIONS:
In this paper, we propose an approach for answering drug queries to support drug prescription. Our focus is on how to obtain and rank answers based on incomplete information and provide personalization. To cope with incomplete and noisy data, we allow both exact and close matches when answering queries. We also present an intuitive approach to display answers to users, which aims to help users to understand the ranked results and possibly refine their queries.

REFERENCE:
22. Sadilek, H. Kautz, J. Bigham, “Finding your friends and following them to where you are,” WSDM, 2012.

www.turkjphysiotherrehabil.org