International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 04 | Apr 2024 www.irjet.net p-ISSN: 2395-0072

Disease Prediction and Doctor Recommendation System

Abhinav Kandalkar[1], Suraj Bute [2], Abhigyan Chhajed[3], Vishnu Zanwar[4]

[1,2,3,4] *B.E. (Computer Science & Engineering), Prof. Ram Meghe Institute of Research and Technology, Badnera-Amravati, Maharashtra, India*

Abstract - *Careful examination of patients' health-related medical records can predict the likelihood of various diseases. In addition, having the expertise of that particular pathologist, if needed, facilitates appropriate and effective diagnosis. This paper presents a new approach that uses data mining techniques, i.e. Naive Bayes classification algorithm for disease prediction recommended by experts predicted heart rate, blood pressure by sensors, and other externally visible symptoms like fever, chills, headache The Naive Bayes algorithm measures those characteristics this and predicts the disease. Additionally, they provide all the necessary and sufficient information about the patient's prognosis and the recommended physicians. The recommendation shows the location of the pathologists, their contacts, and other required information based on the user's chosen filter from low cost, more experience, nearby locations, and they use Stanford's The core NLP algorithm compares physician claims with survey research. Thus, users can access appropriate treatment and necessary medical advice as soon as possible. Additionally, users provide their feedback to the recommending physicians who then incorporate it into search results to provide other recommendations based on the search results.*

Keywords: Core NLP, Naive Bayes, Prediction, Recommendation, Django.

1. INTRODUCTION

Healthcare indeed produces a vast amount of data yearly; this includes medical documents that contain incredible information about the patients' health conditions. However, this task might not only be challenging but also critical due to diagnosis and disease prediction in healthcare. Analyzing all these huge data sets allow us to recognize these patterns, trends, and risks associated with a number of conditions. Through the use of advanced analytics, machine learning algorithms, and AI-based resolutions, we can predict the advent of particular diseases with increasing accuracy. Through this type of forecasting, one can achieve the goal of health protection for the society by way of preventive methods and intervention. An intelligent system for the disease prediction is of great importance for the disease control and the maintaining of health status among the people. It enables the accurate and reliable disease risk prediction based on the accumulated data analytics which in turn helps healthcare professionals to make informed decisions and direct targeted interventions. To further advance healthcare personalization, the system makes it possible to recommend and prescribe treatment according to risk profiles and patient-specific needs. This approach not only leads to better patient outcomes, but enables a more efficient use of healthcare resources and reduces healthcare costs in the long run. In brief, intelligent systems integration in healthcare will make disease control easier to achieve, wellness promotion more effective and healthcare delivery more efficient.

1.1 Prediction System

The main objective of this manuscript is to mine the data and explore the inner coded structures behind the symptom-disease correlations. The disease diagnosis process involves examination of medical profiles based on key parameters like blood sugar levels, blood pressure readings, blood oxygen saturation, presence of headaches, and other symptoms possible in the context condition. An important method for disease classification which is applied is the Naive Bays classifier that is a probability model which the likelihood of a disease is calculated by different observed symptoms and the medical records. The classifier is used to enable the system to identify the seeming complication that primarily affects an individual. The combination of data mining methods with the advanced classifiers such as Naive Bayes greatly increases the disease prediction models' accuracy and trustworthiness. It permits a more subtle comprehension of the dynamics between diseases, symptoms and background factors, helping medical professionals to make more precise diagnoses and treatment decisions. In addition, this technique helps in identifying early warning signs and risk factors to that proactive interventions being taken with preventive measures that would ensure people attain optimal health. However, a variety of data mining programmes and methods of prediction together make a real difference in disease prediction and healthcare.

1.2 Recommendation system

In addition to disease operations through data mining and the use of Naive Bayes Classifier, the system gives a professional on demand via the selected filters by a user. The review-oriented recommendations are obtained by collecting and processing reviews about various physician from other users. These reviews are then processed using the CoreNLP for sentiment analysis and more valuable information extraction. The system seeks the closest experts to the user's location in a case where location-based

recommendations are required, thus providing convenience and accessibility. Payment based and experience based advices are based on the criteria of the fee range and expertise level of medics among specialists. These customizable recommendations improve user interaction with the healthcare system, contributing that they make educated choices about their medical care.

2. MOTIVATION

As of now sensitivity of body to diseases requires regular analysis of body. One of the most significant changes that has occurred today is the availability of multiple platforms for both individual insight prediction and recommendation purposes but the future is likely to be found in composite models that consist of the two systems. Moreover, I think that it will be quite easy and handier if patients could get some primary diagnoses and consultations online - not in hospitals and clinics only. This is done to make sure that consumers save money and spend less time. When some strange results are found in the diagnosis then potential suggestions of nearby specialists and hospitals from the users' list will help to get right and fast curing. Healthcare, becoming a changeable discipline, generating massive data, requires the data to be transcoded into useful knowledge, which can be the reason why large organizations highly invest in this area.

3. RELATED WORK

Binal T. et al [1], Healthcare decision support system for swine flu prediction using naive bayes classifier, This article aims at discovering the medical diagnoses of swine flu by creating a naive Bayes classifier. The classifier categorizes patients into three groups: less likely, least likely, probable, and most likely swine flu cases. The approval rate of the study was around 63.33%, which is moderate predictive power. An important problem mentioned was narrow dataset size leading to lower generalizability and robustness of the model. Shengyong, et al [2], Predicting Disease by using data mining based on healthcare information system, The research team in this study focuses on disease prediction through data mining, especially for cardiovascular diseases. The study compares the performance of three algorithms: Naive Bayes, J-48 decision tree, and a J-48 classifier ensemble with 5 J-48 classifiers. The outcomes show that both Naive Bayes and J-48 classifiers yielded similar accuracy rates of approximately 83%. This implies that these algorithms perform the function of predicting hypertension from medical records in a correct way.

In the following paper, they postulate a collaborative filtering approach that exploits user review data to enrich item descriptions. Core NLP is employed with reviews; then, an algorithm builds representations for computing item similarity using a collaborative filtering framework based on k-nearest neighbors. This way of doing things helps recommendation systems to take the user feedback and preferences into consideration and thus improving the accuracy and relevance of recommendations.

Manjusha K. and her colleagues [5], in Prediction of different dermatological conditions using Naive Bayesian classification, This work is done to predict dermatological conditions using Naive Bayesian classifier. It looks at the probabilities of eight various diseases based on patient attributes and acquires knowledge that is hidden in the database to improve the process of predictions. The system further demonstrates the possibility of predicting diseases other than the dermatological ones, illustrating the universal and adaptable nature of the classification model.

4. PROPOSED WORK

Artificial intelligence makes disease prediction and medical recommendation more convenient and effective by allowing users to fact in the symptoms and real-time vital sign analysis simultaneously. This framework guarantees sufficient accuracy of diagnosis and increases the best possible therapy for patients. The system becomes more user-friendly by integrating user-specific features like distance, rate, assessment, and reviewing.

The system is equipped with a sensing module that is constantly aware of pulse rate, blood pressure, and sugar levels so as to be able to give instant health condition updates. This data, which when integrated with the external symptoms reported by people, underwent a detailed observation process to help in an accurate disease prediction. Advanced algorithms are employed by the system to utilize priorities set by the user in order to locate the most appropriate healthcare organizations.

Users get the liberty to decide how the factors like the proximity, affordability, experience and quality of care would weigh when they are getting the recommendations for the hospitals or clinics. Besides, the system takes doctors' specialties into account, which means patients can get the help of doctors with the skills needed for curing their illnesses personally. These personalized measures further raises the chance for receiving unique medical care and right medication.

Moreover, it can be enriched by the reviews the users will provide on chosen physicians. The feedback mechanism therefore improves the system's precision of recommendation and relevance in a step-by-step fashion, since it adjusts to each patient's changing health care situation and demands.

In total, this integrated approach enables not only precise disease diagnostics detection but also simplification of the access to the top-quality healthcare institutions. Through the provision of information that is patient-centric and the creation of linkages between patients and the most qualified health providers, the system improves patient’s health outcomes and the healthcare system’s efficiency.

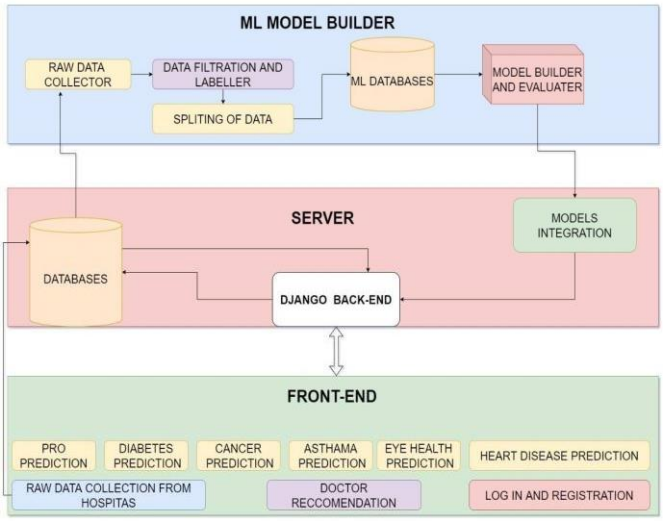

Figure 1: System Overview Diagram

5. IMPLEMENTATION TECHNIQUES

5.1 Naive Bayes

A Naive Bayesian classifier is a superior model presenting the common probability distribution over a stochastic variable set. For a two- class problem, the cases are what's presented to the classifier in which the point variables have different values. The classifier coming computes a posterior probability distribution over the class variable that estimates the probability of each class taking the input features. The literacy of a Naive Bayesian classifier we regard as a construction of the previous chances of the colorful classes, these being the original hypotheticals for the liability of each class appearing. likewise, the predictor calculates tentative chances of the factors given to all of the classes. The tentative chances embody the probability of entering concrete point values given a specific class. The Naive Bayesian classifier accomplishes this through the admixture of previous chances and tentative chances. The classifier is so effective at landing probabilistic connections between features and classes that's suitable to prognosticate what the class marker of a new case is. This system is veritably pivotal and thus applicable in circumstances where the attributes are innately independent or in situations where the supposition of the trait interdependence is reasonable. The model is suitable to perform inversely well indeed with limited training data.

According to Bayes theorem of probability theory:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \quad (1)$$

It is commonly assumed that attributes E1 through Em are conditionally independent given the class label, which means it is often assumed that

$$P(E|H) = \sum_{i=1}^{N} P(Ei = vi|H) \quad (2)$$

After making the above assumption, the classifier is called Naive Bayes classifier.

Table 1: Input Attributes

Gender: F/M

Blood Sugar: 72-162 mg/dl

Blood Pressure: 120/80 mmHg

Palpitation Rate: 60-100 bpm (normal)

Fever: Yes/No

Level of Fever: low(1), moderate(2), high(3)

Cough, Cold: <4 days =1, >4 days =2

Pain: eyes, head, casket, muscles, body, abdominal, no pain

Breathlessness: Yes/No

Headache: normal=1, severe=2

Nausea and vomiting: Present/absent

Feeling weak: Yes/no

Experiencing chills: Yes/no

Constipation: Yes/no

Dizziness: Yes/no

Loss of appetite: Yes/no

the labors being either positive or negative bones , the classifier's performance can be assessed exhaustively, and the classifier's strengths and sins in the bracket of cases can be revealed. Accordingly, Django's libraries go beyond the simplification of the developer's job since they give superior tools for assessing the model's delicacy, performance measures, and confusion matrix that enable data scientists and judges to reach sound opinions after interpreting the results.

6. RESULTS

The illness that the model foresees will be the one that enjoys the highest probability calculated from the Naive Bayes classifier, which is expected to have a 80 percent expected accuracy or higher. Experts that are the best will be recommended within the top five doctors associated with the expected disease and filters selected by the

user. Building on the foundation of the Naive Bayes algorithm, we will proceed with predicting the most likely disease with the expected accuracy of over 80%. This will be followed by providing a list of leading physicians who are experts in treating that foreseen disease. It entails the employment of health specialists' expertise and their nodding with the specific user needs as well as the user's preferences. The Naive Bayes classifier is a probabilistic machine learning method for a variety of tasks including forecasting diseases conditioned on the symptoms and patients profiles. It analyzes the give data and determines the possible diseases by the given data. Then, it selects the disease with the highest probability among them as the predicted disease. An accurate model, with at least N accuracy percentage, demonstrates a reliable prediction methodology, which however, should be open to refining for best performance. The next step is to name the disease for which the model offers the top five specialists with expertise in that condition. The normally tailored recommendation based on the user selected filters might include a number of factors such as the location , specialization ,experience , patient reviews and other preferences. The processing the system puts through filters ensures that recommended professionals do not only have high-end qualifications but also do meet the set requirements of users as well as their expectations.

7. CONCLUSION AND FUTURE WORK

Massive databases of real cases from medical records with assistance from CoreNLP(Natural Language Processing) doctors' recommendation(s) are the base of such algorithms. The introduction of more symptoms as well as additional cases for both training and testing purposes will not only contribute to the precision but also make the models more effective.

Beyond this, integrating different health metrics including calorie intake, step count, sleep quality and other necessary medical profiles reportable from wearable devices can further avail a complete status of an individual's health. Using this plentiful of data, nutritional plans can be prepared and designed individually for each user according to their unique needs and health goals.In addition, to the existing measures, the real-life alerts and notifications can be triggered. When certain health risks or emergencies are detected, the user and designated caregivers or parents received the timely alerts immediately. This insures the timely intervention hence better management of health and preventative care treatments.

Through continuous improvement and enlargement of these systems integration, predictive healthcare could transform entirely how diseases are diagnosed, managed, or prevented. The synergistic effect of data analytics, machine learning, natural language processing, and wearable technology leads to a holistic and preventative approach to healthcare delivery allowing users to take a proactive position on their health while strengthening the collaboration among patients, healthcare providers, and caregivers.

References

[1] Binal A. Thakkar, Mosin I. Hasan, Mansi A. Desai, “Healthcare decision support system for swine flu prediction using Naive Bayes classifier”, “IEEE”, 101-105, 2010.

[2] Feixiang Huang, Shengyong Wang, Chien Chung Chan, “Predicting disease by using data mining based on healthcare information system”, “IEEE”, 2012.

[3] Rafael M. D’addio, Marcelo G. Manzato, “A collaborative filtering approach based on users’ reviews”,“IEEE”, 204209, 2014.

[4] F. O. Isinkaye, Y.O. Fola Jimi, B.A. Ojokoh,“ Recmmendation systems : Principles, Methods and Evaluation”, “Elsevier”, 261-273,2015.

[5] Manjusha K.K.,K. Sankaranarayanan, Seena P,“Prediction of different dermatological conditions using Naive Bayesian classification”,”IJARCSSE”,864-868,2014.

[6] Subhash C. Pandey, “Data Mining techniques for medical data: A Review”, “IEEE”, 2016.

[7] Lee, Yunkyoung, “Recommendation system using collaborative filtering”, 2015.

[8] Geoff A., Elvis C., Mickey E., Sameul S., “Smart wearable body sensors for patients self assistant and monitoring”, BioMedCentral, 2014.

[9] G Gannu, Y Kakodkar, A Marian, “Improving the quality of predictions using textual information in online users reviews”, Information Systems, 2013.

[10] Sellappan P., Rafia A., “Intelligent heart disease prediction system using data mining techniques”, IEEE/ACS, 2008.