A LOCALISED MODEL FOR DOCTOR RECOMMENDATION ON CARDIOLOGIST USING NAIVE BAYES

Abstract

The health-world industry has seen a huge transition from the traditional healthcare system to the AI- based healthcare sector. However, the lack of a unified method for cost review prior to hospital visits means that overall costs for patient treatment remain high and difficult to manage. The ease with which patients may access modern technologies has contributed to the growing popularity of online medical systems in recent years. A method for early illness prediction based on symptoms is what we suggest. This technology using machine learning also schedules an appointment based on doctor availability and recommends hospitals in the area based on user ratings and facility locations. By automating the illness prediction, this system aims to decrease the time and effort spent by patients and clinicians.

Keywords: Machine Learning, Disease Prediction, Doctor Recommendation, Personalized Healthcare, Healthcare Delivery, Algorithmic Approaches

1. Introduction

Every year, the healthcare sector produces gigabytes of data.
The archived medical records contain a wealth of patient-related data. Retrieving pertinent data or high-quality medical care is a difficult but crucial endeavour [1]. Through the analysis of these vast amounts of data, we are able to forecast when the disease will strike and protect individuals. Thus, by offering precise and reliable illness risk prediction, an intelligent system for disease prediction contributes significantly to the management of disease and the preservation of people's good health.

1.1 Objective

A smart healthcare system has following main objectives:

a. Develop a Nave Bayes classification model that identifies the illness according to user comments.
b. To construct a framework for a web interface for illness prediction.

c. To recommend localized specialist for the specific cardio disease

1.2 Causes of Cardiovascular Disease

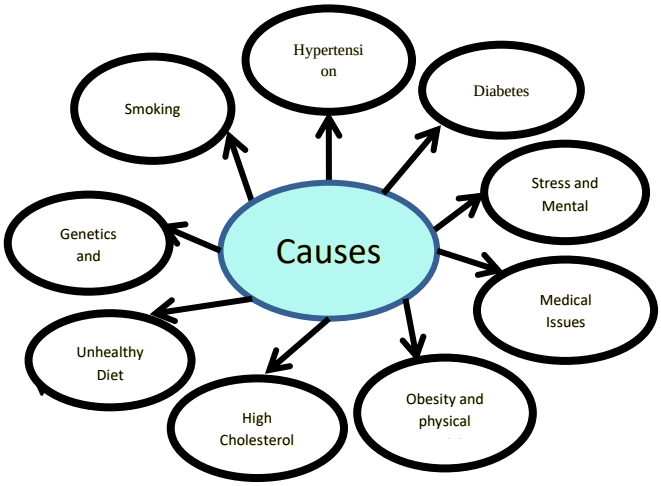

Fig.1. Cardiovascular Disease Causes

1. Related Work

This study is the result of our examination of the below publications from various additional sources. We looked at the link between various algorithms and how well they performed for illness prediction in various settings.

Binal T. et al [2], Healthcare decision support system for swine flu prediction using naïve bayes classifier, focuses on the aspect of medical diagnosis by learning patterns through the collected data for swine flu using naïve bayes classifier for classifying the patients of swine flu into three categories (least possible, probable or most probable), resulting into an accuracy of nearly 63.33%. Datasets used for this classification were limited in number.

In their paper "Predicting Disease by using data mining based on healthcare information system," Shengyong W. et al. [7] detail the trials conducted on data mining for the purpose of disease prediction using a substantial amount of actual hypertension medical records. In this essay, three algorithms: ensemble of five J-48 classifiers, J-48, and naïve Bayes. The accuracy of 83% was almost the same for J-48 and Naïve Bayes in this case.

Marcelo M. et al [3], A collaborative filtering method based on user reviews suggests a method that generates item descriptions by utilising user reviews. After processing the reviews with the CoreNLP tool, an algorithm is created to calculate the similarity of the items, which is then employed in a collaborative filtering technique based on k-nearest neighbours.

In their paper Prediction of diverse dermatological problems using Naive Bayesian classification, Manjusha K. et al. [8] list eight possible diseases based on the patient's characteristics. The database's hidden knowledge is extracted by the system. In addition to dermatological illnesses, the system can anticipate other diseases.

2. Main Contribution

A machine learning (ML) doctor recommended system's primary benefits usually centre on optimising the distribution of medical resources, boosting patient experience, and improving healthcare services. The following are some significant contributions:

- Data analysis is essential in the data-driven world of today. It aids businesses in utilising data to make choices, streamline operations, and obtain a competitive advantage. By transforming unstructured data into valuable insights, data analysis enables businesses to recognize opportunities, reduce risks, and improve overall performance.

- The Naive Bayes algorithm is a supervised learning technique that solves classification issues. It is based on the Bayes theorem.
  Its primary application is in text categorization, where a high-dimensional training dataset is used.
  One of the most straightforward and efficient classification algorithms, the Naive Bayes classifier aids in the rapid development of machine learning models with rapid prediction capabilities.
  Being a probabilistic classifier, it makes predictions based on the likelihood that an object will occur.

- A collection of Java-written natural language analysis tools for text processing is offered by CoreNLP. It creates a comprehensive structured study of the most popular NLP processes from raw English text input. Annotators are a collection of language analysis tools that are integrated into CoreNLP [3]. Sentence splitting, tokenization, part-of-speech (POS) tagging, and sentiment analysis are a few of the pertinent annotators included in this programme. CoreNLP methods for physician suggestion based on prior user reviews have been researched and put into practice.
  By adding additional symptoms and increasing the quantity of cases for testing and training, the system can be made even better. A required nutritional diet plan can also be recommended by taking into account the number of calories, steps taken, sleep quality, and other medical profiles through wearable technology. In the event of a risk, timely alerts and notifications can be delivered to the user's guardian as well.

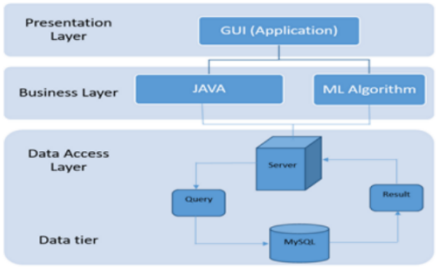

Fig2. Architecture Diagram

- A doctor recommendation system's user interface (UI) is designed to be user-friendly and intuitive, enabling patients to find suitable healthcare professionals based on their specific needs. Search Results where physicians who meet the search parameters are listed on the search results page. A brief biography, ratings, the doctor's name, and specialization may be included in each entry. To further refine the results based on variables like years of experience, years of proximity, availability, and languages spoken, filters may also be available. Physician Profile where selecting a physician's name may direct the user to a website featuring additional information about that physician. This might have a button for making an appointment as well as information on their training, credentials, and patient testimonials. Appointment Booking where users would be able to select a time and day for their appointment, provide personal details, and explain any symptoms or reasons they are visiting. Review and Ratings where users will be asked to rate their experiences.

3. Proposed Methodology

There are various techniques for illness prognosis. Nonetheless, research has been done on heart-related conditions, and a risk level

has been computed. These techniques aren't typically applied to illness prediction, though. Thus, general disease prediction is aided by the smart healthcare system. Sometimes, you or a family member may require immediate medical attention.

Naive Bayes:
The suggested smart healthcare system framework uses a data mining approach called the "Naive Bayes classifier" to develop the expectation framework. [4] In order to ascertain the precise symptom expectation, this system has a higher quantity of data indexes and features that are actually acquired from expert data. A number of artificial intelligence and data mining methods rely on the "Naive Bayes or Bayes" Rule.

Precognitive ability models are created using the norm. By establishing the connection between the objective (i.e., subordinate) and other variables, it gains from the "proof".

Model Diagram of Naive Bayes

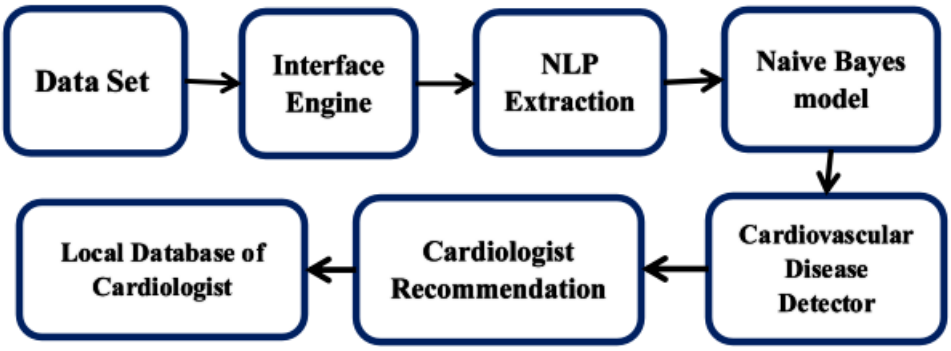

Fig.3. Model Diagram

A Naive Bayesian classifier, is a model joint probability distribution over a set of stochastic variables. Instances of the classification problem under study are presented to the classifier as a combination of values for the feature variables;

the classifier then returns a posterior probability distribution over the class variable. Learning such a classifier amounts to establishing the prior probabilities of the different classes and estimating the conditional probabilities of the various features given each of the classes [2]. According to Bayes theorem of probability theory:- (1) It is assumed that attributes E1 to Em are class conditionally independent, which means it is often assumed that

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

Where,

P(H|E) is Posterior probability: Probability of hypothesis H on the observed event E.

P(E|H) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P(H) is Prior Probability: Probability of hypothesis before observing the evidence.

P(E) is Marginal Probability: Probability of Evidence.

Data Flow Diagram

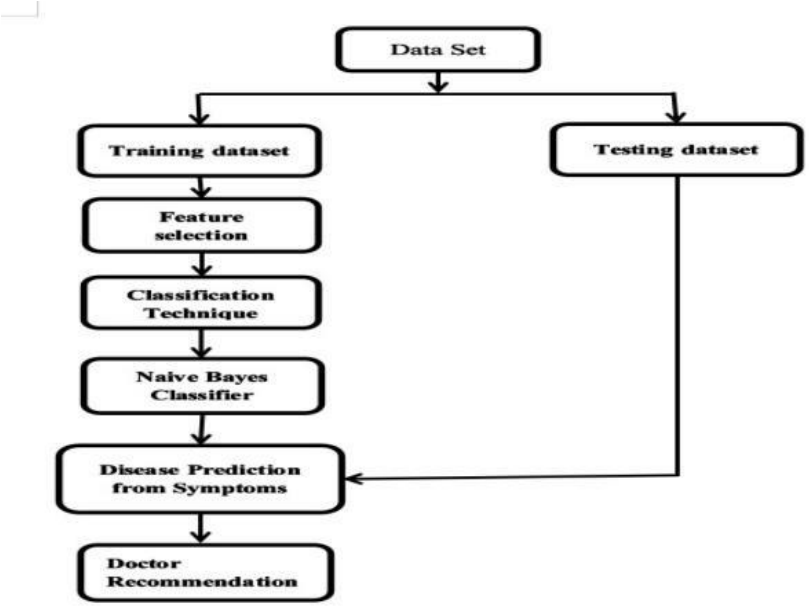

Fig.4. Data flow diagram of the process followed using Naive Bayes

4. Result and Analysis

The Naive Bayes classifier's probability calculations in fig. 5 indicate the most likely sickness, with an expected accuracy of more than 90%. In fig.6 there is a distinguished graph that shows the representation of top five specialists recommended based on user-selected filters and the anticipated disease. The statistical result according to other algorithms with respect to Naive Bayes is as follows:

ALGORITHMS USED	% OF ACCURACY
Naive Bayes	96.2
Data Pre-processing	90.1
Matrix Factorization	85.3
Recommendation Model	88.8
Singular Value Decomposition(SVD)	81.2

Fig. 5. Graph of result analysis

The graph for the following result is :

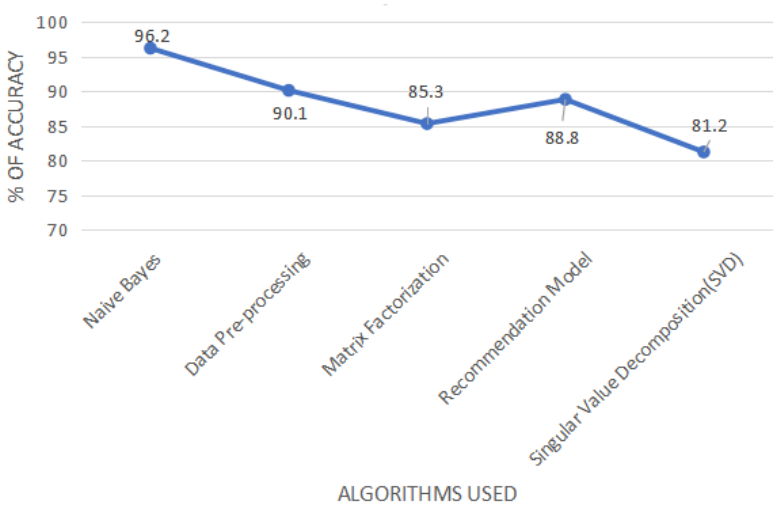

Fig. 6. Distinguished Graph of Naive Bayes Classification

References:

1. Subhash C. Pandey, “Data Mining techniques for medical data: A Review”, “IEEE”, 2016.
2. Binal A. Thakkar, Mosin I. Hasan, Mansi A. Desai, “Healthcare decision support system for swine flu prediction using naïve bayes classifier”, “IEEE”, 101-105, 2010.
3. Rafael M. D’addio, Marcelo G. Manzato, “A collaborative filtering approach based on users’ reviews”,“IEEE”, 204- 209, 2014.
4. T. Devasia, T. P. Vinushree and V. Hegde, "Prediction of students performance usingEducational Data Mining", Proceedings of IEEE International Conference on Data Mining and Advanced Computing (SAPIENCE), pp. 50-100, 2016
5. Sellappan P., Rafia A., “Intelligent heart disease prediction system using data mining techniques”, IEEE/ACS, 2008.
6. G Gannu, Y Kakodkar, A Marian, “Improving the quality of predictions using textual information in online users reviews”, Information Systems, 2013.
7. Feixiang Huang, Shengyong Wang, Chien Chung Chan, “Predicting disease by using data mining based on healthcare information system”, “IEEE”, 2012
8. Manjusha K.K.,K. Sankaranarayanan, Seena P,“Prediction of different dermatological conditions using naïve Bayesian classification”,”IJARCSSE”,864-868,2014.