Smart Healthcare System

Malavika Padmakumar Nair
Computer Engineering
Vidyavardhini's College of Engineering and Technology
Vasai, India

Pooja Rajubhai Nayaka
Computer Engineering
Vidyavardhini's College of Engineering and Technology
Palghar, India

Jinisha Gangadharan Karuvath
Computer Engineering
Vidyavardhini's College of Engineering and Technology
Vasai, India

Smita Jawale
Computer Engineering
Vidyavardhini's College of Engineering and Technology
Vasai, India

***Abstract*— Classical diagnosis is a process in which a medical practitioner examines the body of a patient for any possible signs or symptoms of a medical condition, undergo various medical examinations, and then conclude. This is often challenging because many signs and symptoms are nonspecific. Medical facilities must be advanced so that better decisions for patient diagnosis and treatment options can be created. Machine learning allows models to rapidly analyze data, and deliver results, leveraging both historical and real-time data. Healthcare service providers can make better decisions on patient's diagnoses and treatment options, which will end up in overall improvement of healthcare services and thereby increasing patient satisfaction. By this project, we are trying to implement functionalities of machine learning in healthcare within a single system. Some cases can occur when an early diagnosis of a disease is not within immediate reach. In such cases, this system can be effectively implemented. As widely said, "Prevention is best than cure", prediction of diseases and epidemic outbreak would lead to early prevention of an occurrence of a disease. This project mainly focuses on the development of a system, or we could say an immediate medical provision that would incorporate the symptoms and other medical data collected from the patient and store them into a Smart health dataset. This dataset would then be analyzed using the Naïve Bayesian machine learning algorithm to deliver results with maximum accuracy. GPS tracking will be used to suggest nearest doctor or specialist if the patient needs referral. This system will have the capacity to reduce costs of treatment, predict outbreaks of epidemics, avoid preventable diseases and improve the quality of life. Patients can seek information regarding diseases, pathological tests referred, get answer to any kind of question and solve any problem related to the disease.**

I. INTRODUCTION

In India more than 70% of people in India are prone to general body diseases like viral, flu, cough, cold, etc. in every 2 months. Because many of us do not realize that the overall body diseases might be symptoms to something more harmful, 25 you look after the population succumbs to death due to ignoring the early general body symptoms. This could be a dangerous situation for the population and can be are alarming. Hence identifying or predicting the disease at the earliest is extremely important to avoid any unwanted casualties. The currently available systems are the systems that are either dedicated to a specific disease or are in research phase for algorithms when it involves generalized disease [1].

II. AIMS AND OBJECTIVES

The purpose of this technique is to supply prediction for the overall and more commonly occurring disease that when unchecked can become fatal disease. The system applies data mining techniques and Naïve Bayesian algorithm. This system will predict the most probable disease supporting the given symptoms and measures required to avoid the hostility of disease. In this project, the system will be trained using machine learning.

III. LITERATURE REVIEW

1. EXISTING SYSTEM

The current system is a manual system where the patient visits a doctor to get diagnosed. Also records of patients and their medical conditions are just maintained on paper. This process is quite time consuming and it is also not possible for everyone to identify their diseases at their home easily. Some sources available as individual prediction or recommendation system are not integrated together and their efficiency in clinical settings is still debatable [2]. Also there exists no system to recommend various pathological tests.

2. PROPOSED SYSTEM

A computer aided system to use intelligent machine learning techniques to predict the most accurate illness based on the inputted symptoms. It also consist of doctor recommendation system, which will help the patient locate a doctor in the nearest locality who is an expert in the domain of disease suggested. The system would also maintain a database of the patients which contains medical details of the patient. This medical expertise system also aims to predict the outbreak of any disease

within a locality and help in taking the necessary actions required. The system aims to reduce the time factor which is being consumed in the existing system, make it available to the common people, especially in areas where the medical facilities are low and rule out the paper-based works.

IV. SCOPE

The proposed method have presented an intelligent and effective prediction methods by using machine learning methods. This proposed method briefly examines the potential use of supervised machine learning algorithms such as Naive Bayes to massive data volume. For pre-processing of data and data prediction naive Bayes categorization are used. Many experimental results has shown that our Naive Bayes classification can achieve more accuracy even when used with more than 50 attributes in our datasets. Ultimately several factors will be evaluated from this prediction results such as recall precision, accuracy, etc. Firstly, this method provided an efficient approach for the extraction of significant patterns from the disease dataset for the efficient guess of various disease. Based on the computation, the numerous patterns of disease having greater value than a predefined threshold will be chosen for the valuable prediction of various diseases. The goals are defined based on intelligence and exploration of data. The goals got to be calculated against the trained models. The system is so designed as to help everyone identify possible diagnosis of early massive disease. This system was thoroughly tested to ensure that it is consistent and valid. It is hoped that the system will go ahead to earlier detection of chronical diseases, lower healthcare costs and, most significantly, better quality of life for people suffering from chronical diseases.

V. FLOWCHART

VI. IMPLEMENTATION METHODOLOGY

Module 1: Disease Prediction

1. Naïve Bayesian algorithm : This system accepts the input from the user and predicts the most probable disease. This is achieved with help of dataset and the machine learning algorithm. The algorithm here is Naive Bayesian which works on probabilistic approach. We have imported Scikit learn library for Naïve Bayesian algorithm. For this we have used Multinomial NB since multiple variants i.e. multiple symptoms are taken. It is a classification algorithm which calculates the probability of any event occurrence.

2 .Multinomial naive bayes:
It estimates the conditional probability of a particular instance given in a class as the relative frequency of term belonging to class. The variation takes into account the number of occurrences of term in training from class , including the multiple occurrences.

3. Bayesian Theorem : The purpose of Bayesian theorem is to predict the class label i.e. disease in our project for a given tuple. Let X be a tuple containing symptoms and H be some hypothesis, such as that the data tuple X (symptoms) belongs to a specified class C (disease). For problems on classification, we look for the probability that tuple X belongs to class C, given that we know the attribute description of X.

4.Dataset : The dataset was taken from study conducted at Colombia University. It consist of 150 diseases and each disease consist an average of 8-10 symptoms. 70% of dataset which used for training was made considering all combinational inputs. The symptoms present for corresponding disease was marked as 1 and remaining 0. It consists 5 drop-down options where we have passed a list of symptoms. The user can select any five symptoms and clicking the predict button the disease predicted will be displayed in the text-box.

Disease	Count of Disease Occurrence	Symptom
UMLS:C0020538_hypertensive disease	3363	UMLS:C0008031_pain chest
		UMLS:C0392680_shortness of breath
		UMLS:C0012833_dizziness
		UMLS:C0004093_asthenia
		UMLS:C0085639_fall
		UMLS:C0039070_syncope
		UMLS:C0042571_vertigo
		UMLS:C0038990_sweat^UMLS:C0700590_sweating increased
		UMLS:C0030252_palpitation
		UMLS:C0027497_nausea
		UMLS:C0002962_angina pectoris
		UMLS:C0430716_pressure chest
UMLS:C0011847_diabetes	1421	UMLS:C0032617_polyuria
		UMLS:C0085602_polydypsia
		UMLS:C0392680_shortness of breath
		UMLS:C0008031_pain chest
		UMLS:C0004093_asthenia
		UMLS:C0027497_nausea
		UMLS:C0085619_orthopnea
		UMLS:C0034642_rale
		UMLS:C0038990_sweat^UMLS:C0700590_sweating increased
		UMLS:C0241526_unresponsiveness
		UMLS:C0856054_mental status changes
		UMLS:C0042571_vertigo
		UMLS:C0042963_vomiting
		UMLS:C0553668_labored breathing

Fig: Data taken

Module 2: Doctor Recommendation

Based on the disease predicted the patient is recommended a doctor in the nearest locality. The doctors recommended for the predicted disease are specialists in that field. A database of doctors is maintained which contains attributes like doctor's name, specialization, location of clinic or dispensary and contact number. With the help of the location attribute the patient is suggested a doctor.

Module 3: Pathological Test Suggestion

Certain chronical diseases require pathological tests to be carried out. So the system would suggest the patient the required test that needs to be carried out for the disease that is predicted. The test names for that particular disease is stored in a database. Each time a disease is predicted, it is fetched in the database and the corresponding tests are suggested.

Module 4: Outbreak Prediction

With the help of GPS patient diagnosed with a disease will be tracked and recorded. If more number of patients in the same locality are predicted with the same disease, then the admin notifies that there is an outbreak of that disease. A notification system resembling a noticeboard will be used to notify regarding this outbreak [3]. The admin will have access to this notification page and others can view the notifications.

VII. ALGORITHM

```
class MultinomialNB(_BaseDiscreteNB):

    Parameters
    ----------
    alpha : float, optional (default=1.0)
        Additive (Laplace/Lidstone) smoothing parameter
        (0 for no smoothing).

    fit_prior : boolean, optional (default=True)
        Whether to learn class prior probabilities or not.
        If false, a uniform prior will be used.

    class_prior : array-like, size (n_classes,), optional (default=None)
        Prior probabilities of the classes. If specified the priors are not
        adjusted according to the data.

    Attributes
    ----------
    class_count_ : array, shape (n_classes,)
        This is number of samples encountered for each class during fitting. The obtained
        value is then weighted by the sample weight when it is to be provided.

    class_log_prior_ : array, shape (n_classes, )
        Smoothed empirical log probability for each class.

    classes_ : array, shape (n_classes,)
        Class labels known to the classifier

    coef_ : array, shape (n_classes, n_features)
        Mirrors ``feature_log_prob_`` for interpreting MultinomialNB
        as a linear model.

    feature_count_ : array, shape (n_classes, n_features)
        This is Number of samples encountered for each (class, feature)
        during fitting. This obtained value is then weighted by the sample weight when it is to be
        provided.

    feature_log_prob_ : array, shape (n_classes, n_features)
        Empirical log probability of features
        given a class, ``P(x_i|y)``.

    intercept_ : array, shape (n_classes, )
        Mirrors ``class_log_prior_`` for interpreting MultinomialNB
        as a linear model.

    n_features_ : int
        Number of features of each sample.
```

VIII. FORMULA

The Naive Bayes method is a classification method based on Bayes' Theorem and the conditional independence assumption.

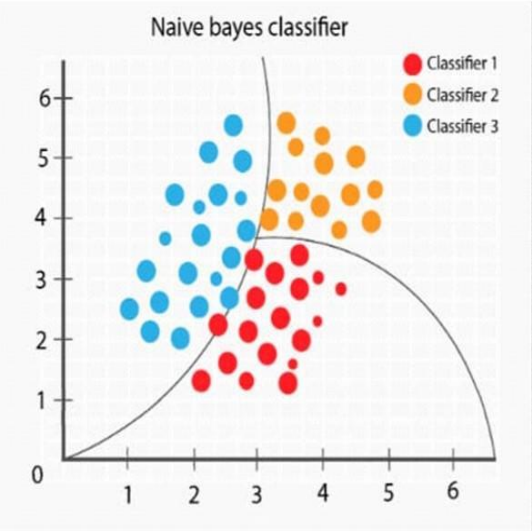
For a given training set,
T = {(x1,y1) , (x2,y2) , (x3,y3), ... , (xn,yn)}.

Naive Bayes first learns the joint probability distribution (X,Y) of the input and output by the conditional probability distribution based on the conditional independence assumption. The output label with the biggest posterior probability for the given input can be calculated.

Bayes theorem [4] says that,

$$p(cj/d) = p(d/cj) * p(cj) / p(d)$$

where,
p(cj/d) is probability of instance d being in class cj.
p(d/cj) is probability of generating instance d given class cj.
p(cj) is probability of occurrence of class cj.
p(d) is probability of instance d occurring.



IX. RESULT

We used supervised machine learning algorithm to predict the most probable disease from the probabilities of diseases calculated by Naïve Bayesian algorithm. After disease prediction, the patient can also get the details of the nearest doctor with the help of GPS tracking. The system suggests the required pathological tests to the patients. The admin can also notify if there is an outbreak of a particular disease.

X. CONCLUSION

A smart system for accurate disease prediction and medical facility recommendation plays a major role in effective treatment. By applying Machine Learning techniques for disease prediction on a large number of real world dataset accurate results have been achieved. In addition to this, the specialists for the predicted disease are recommended based on the location that the user selects. Disease based pathological test suggestions are also done based on the disease that is predicted.

XI. REFERENCES

[1] Allen Daniel Sunny, Sajal Kulshreshtha, Satyam Singh, Srinabh, Mr. Mohan Ba, Dr. Sarojadevi H. , "Disease Diagnosis System By Exploring Machine Learning Algorithms", in International Journal of Innovations in Engineering and Technology(IJIET)

[2] Dhanashri Gujar, Rashmi Biyani, Tejaswini Bramhane, Snehal Bhosale, Tejaswita .P. Vaidya, "Disease Prediction and Doctor Recommendation System", in International Research Journal of Engineering and Technology(IRJET).

[3] Konde T.R., Konde D.R., Khokrale P.V, Phulwade S.P, "Health Prediction System" in International Journal of Advance Engineering and Research Development (IJAERD), Technophilia – 2018.

[4] Han, Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann 3rd edition. .