Data Driven Herb's Product Recommendation Framework

Aishwarya Patil , Shubhangi Chavhan , Vaishnavi Chimate, Apurva Chinni ,

Asst. Prof. N.A.Mulla

B.E. Students, Assistant Professor, Department of Information Technology Engineering, Bharati Vidyapeeth's College of Engineering for Women, Pune, India.

Abstract:

In this era, the burden of chronic disease is rapidly increasing worldwide. Also varieties of disease are on rise. This eventually makes the prediction of disease extremely vital. The main ideology is to predict the disease and recommend the medication by analyzing the symptoms provided by the user as an input. As Ensemble method improves the result by combining multiple models, this makes us use K-Nearest Neighbor, Naïve Bayes and Random Forest for disease prediction. This also focuses on the recommendation of Herbs' products in accordance with the predicted disease.

Keyword:

Data Mining, Naïve Bayes, K-Nearest Neighbor (KNN), Random Forest,

Introduction:

In recent scenario of medical field, data mining has reduced the complexity for disease prediction. From the gigantic data-sets or other repositories excavating information through data driven process is known as Data Mining . The use of classification algorithms to predict diseases is the main purpose of this concept. The algorithms used are KNN, Random forest and Naïve Bayes. Based on performance factors like classification accuracy and retrieval time, these classifier algorithms are evaluated. The main motive behind selecting these supervised algorithms mostly include the least error rate, easy to implement, high accuracy, less execution time as well as they prove to be the most reliable. In order to generate one optimal predictive model, a technique that fuses several base models is known as Ensemble method. To improve the end result, incorporation of the KNN, Naïve Bayes and Random forest algorithms is preferred.

Mostly plant originated medicines obtained from contrasting composition of plant parts e.g. roots, leaves, or flowers are gentle for people of all age. Herbs, which are considered to be the most reliable as well as the safest medication for well-being of human, are preferred to be recommended. It can be proved beneficial for the novice users to acquaint with information of relevant Herbs' products.

Related work:

[1] 2017, Wei Li, Zheng Yang, "Distributed Representation for Traditional Chinese Medicine Herb via Deep Learning Models "

Description- In this paper, they have proposed the representation of Traditional Chinese Medicine (TCM) herb with distributed representation via Prescription Level Language Modeling. In this experiment, they have calculated the correlation between the medicine and the judgment of professionals by Spearman's Score which gives a result of 55.35

[2] 2016, M. A. Jabbar, B. L. Deeksh Atulu and Priti Chandra, "Intelligent heart disease prediction system using random forest and evolutionary approach "

Description- In this research paper, using Random forest they have developed an effective approach for Heart disease. Using chi square they have adopted feature selection and genetic algorithm measures for heart disease classification. This technique has achieved an accuracy of 83.70%

 [3] 2013, S. Vijiyarani, S. Sudha, "Disease Prediction in Data Mining Technique – A Survey"

Description- In this paper, the main motive was on using various algorithms and blend of several target attributes for different types of disease prediction using data mining. The algorithms used were KNN, Naive Bayes and Decision List. Next they have discussed the feature subset selection using genetic algorithm. The data was divided into 2 parts, 70% was provided for training and 30% for testing. Using 10 fold cross validation the classified data is evaluated.

[4] 2016, Dr. B. Shrinivasan, K. Pavya, "A study on data mining prediction technique in healthcare sector"

Description- For good decision making this paper explores various data mining techniques which are used in medical field. Prediction techniques used are Bayesian Classifier, Artificial Neural Network (ANN), Bayesian Classifier and Decision tree .

[5] 2017, Narender Kumar, Sabita Khatri, "Implementing WEKA for medical xata classification and early disease prediction"

Description- In this paper, authors have compared SVM, J48, Random Forest, Naïve Bayes and KNN classifiers using performance measures like Kappa statistics, R, Mean Absolute Error and Root Mean Square Error using WEKA tool. On various accuracy measures like Recall, TP rate, FP rate, Precision and f –measures, they have compared these classifiers by implementing on WEKA.

[6] 2018, Vineetha S., Sweetlin S., Vinusha H. and Sajini S., "Disease prediction using Machine learning over big data"

Description- In this paper, by using structured as well as unstructured data from hospital a machine learning decision tree map algorithm is developed. They have also used map reduce algorithm for partitioning the data. The screening accuracy of the algorithm is 94.8% with regular speed.
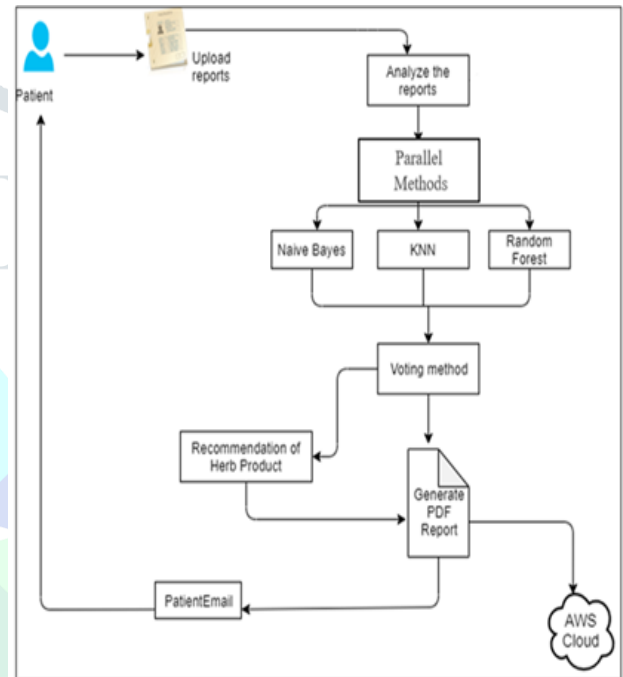
System Architecture:

Fig. Architecture Diagram

The architecture depicts the flow of disease prediction as well as the recommendation framework. This concept contains two major phases viz. prediction phase and recommendation phase. The prediction phase enables the acceptance of the symptoms. The acknowledged data is further examined methodically. Based upon the results gathered, a balloting method is carried out and an appropriate disease is forecasted. In accordance with the disease declared, pertinent herbs' products are suggested. Further a précised PDF report is generated and sent to the user's mail for future invocations. For further references, the detailed PDF is accumulated on the AWS cloud.

Conclusion:

This paper aims on predicting the disease on the basis of the symptoms using an ensemble method and recommend the relevant herbs' product for the forecasted disease. The plan is to design a system in such a way that it accepts symptoms from the user as input and produces output i.e. predicted disease and recommend suitable herbs' products in the most accurate way.

Reference:

[1] Wei Li, Zheng Yang, "Distributed Representation for Traditional Chinese Medicine Herb via Deep Learning Models", 2017.

[2] M .A. Jabbar, B. L. Deekshatulu and Priti Chandra,"Intelligent heart disease prediction system using random forest and evolutionary approach", 2016

[3] S. Vijiyarani, S. Sudha "Disease Prediction in Data Mining Technique – A Survey", 2013

[4] Dr. B. Shrinivasan, K. Pavya, "A study on data mining prediction technique in healthcare sector" , 2016,

[5] Narender Kumar, Sabita Khatri, "Implementing WEKA for medical xata classification and early disease prediction", 2017,

[6] Vineetha S., Sweetlin S., Vinusha H. and Sajini S., "Disease prediction using Machine learning over big data", 2018