

MULTIPLE DISEASE PREDICTION SYSTEM USING MACHINE LEARNING

Nusrat Mohammed Kaleem¹, Sarah², Karim Lakhani³, Dr. T PremChander⁴

UG Scholar^{1,2,3}, Associate Professor⁴

Department of Information Technology, ISL Engineering College, Hyderabad, Telangana.

Abstract - The rapid proliferation of Internet technology and handled devices has opened up new avenues for an online healthcare system. There are instances where online medical help or healthcare advice is easier and faster to grasp than real-world help. People often feel reluctant to go to hospital or physician or minor symptoms. However, in many cases, these minor symptoms may trigger major health hazards. As online health advice is easily reachable, it can be a great head start for users. Moreover, existing online health care systems suffer from a lack of reliability and accuracy. This system analyzes the symptoms provided by the user as input and gives the disease as an output. Prediction is done by implementing the Naive Bayes Classifier.

Keywords - Disease prediction · machine learning · symptoms

I. INTRODUCTION

At present, when one suffers from particular disease, then the person has to visit to doctor which is time consuming and costly too. Also if the user is out of reach of doctor and hospitals it may be difficult for the user as the disease can not be identified. So, if the above process can be completed using a automated program which can save time as well as money, it could be easier to the patient which can make the process easier. There are other Heart related Disease Prediction System using data mining techniques that analyzes the risk level of the patient.

Disease Predictor is a web based application that predicts the disease of the user with respect to the symptoms given by the user. Disease Prediction system has data sets collected from different health related sites. With the help of Disease Predictor the user will be able to know the probability of the disease with the given symptoms.

As the use of internet is growing every day, people are always curious to know different new things. People always try to refer to the internet if any problem arises. People have access to internet than hospitals and doctors. People do not have immediate option when they suffer with particular disease. So, this system can be helpful to the people as they have access to internet 24 hours.

Medicine and healthcare are some of the most crucial parts of the economy and human life. There is a tremendous amount of change in the world we are living in now and the world that existed a few weeks back. Everything has turned gruesome and divergent. In this situation, where everything has turned virtual, the doctors nurses are putting up maximum efforts to save people's lives even if they have to danger their own. There are also some remote villages which lack medical facilities. Virtual doctors are board-certified doctors who choose to practice online via video and phone appointments, rather than in-person appointments but this is not possible in the case of emergency. Machines are always considered better than humans as, without any human error, they can perform tasks more efficiently and with a consistent level of accuracy. A disease predictor can be called a virtual doctor, which can predict the disease of any patient without any human error. Also, in conditions like COVID-19 and EBOLA, a disease predictor can be a blessing as it can identify a human's disease without any physical contact. Some models of virtual doctors do exist, but they do not comprise the required level of accuracy as all the parameters required are not being considered. The primary goal was to develop numerous models to define which one of them provides the most accurate predictions. While ML projects vary in scale and complexity, their general structure is the same. Several rule-based techniques were drawn from machine learning to recall the development and deployment of the predictive model. Several models were initiated by using various machine learning (ML) algorithms that collected raw data and then bifurcated it according to gender, age group, and symptoms.

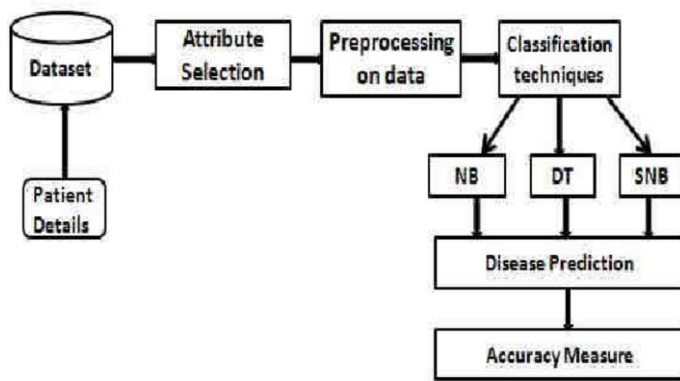
II. PROPOSED SYSTEM

Proposed system for disease prediction. The doctor may not be available always when needed. But, in the modern time scenario, according to necessity one can always use this prediction system anytime. The symptoms of an individual along with the age and gender can be given to the ML model to further process. After preliminary processing of the data, the ML model uses the current input, trains and tests the algorithm resulting in the predicted disease.

III. PROBLEM STATEMENT

The classical diagnosis method is a process where the patient has to visit a doctor, undergo various medical tests, and then come to a conclusion. This process is very time-consuming. To save time required for the initial process of diagnosis symptoms, this project proposes an automated disease prediction system that relies on user input. The system takes input from the user and provides a list of probable diseases. There are many tools related to disease prediction. But particularly heart related diseases have been analyzed and risk level is generated. But generally there are no such tools that are used for prediction of general diseases. So Disease Predictor helps for the prediction of the general diseases.

Block diagram



Module description

The system will predict the disease where the symptoms are given as the input. The disease will be predicted using the Naive Bayesian algorithm. According to the literature survey, this algorithm results in the maximum accuracy for a larger dataset. The dataset contains disease as labels and for each disease, symptoms are given. 70% of the dataset will be used as training and 30% will be used for testing data. Training and testing would be done on the dataset and the desired output will be obtained.

Naive Bayes algorithm

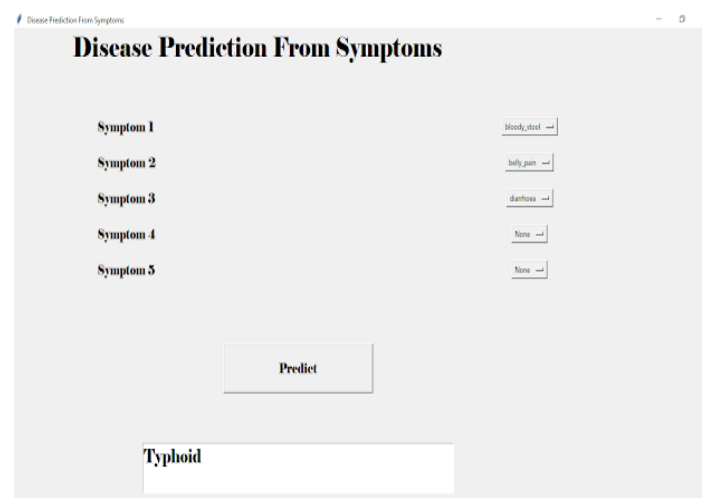
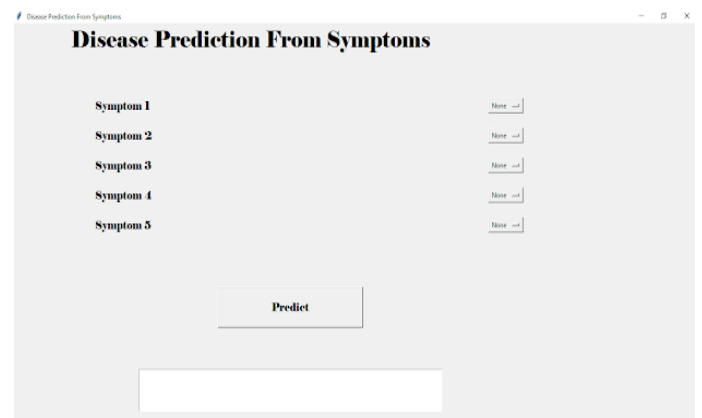
This system accepts the input from the user and predicts the most probable disease. This is achieved with the help of the dataset and the machine learning algorithm. The algorithm here is [Naive Bayesian](#) which works on a probabilistic approach. We have imported [Scikit](#) to learn the library for its implementation. For this, we have used multinomial NB since multiple variants i.e. multiple symptoms are taken.

Bayesian Theorem

- The purpose of the 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 classification problems, we are looking for the probability that tuple X belongs to class C , given that we know the attribute description of X .

Dataset

The dataset was taken from a study conducted at Colombia University. It consists of 150 diseases and each disease consist of an average of 8-10 symptoms. 70% of the dataset used for training was made considering all combinational inputs. The symptoms present for the corresponding disease were marked as 1 and remaining as 0. It consists of 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.



IV. RESULT AND CONCLUSION

The project is designed in such a way that the system takes symptoms from the user as input and produces output i.e. predict disease. The user can select a minimum of one to a maximum of five symptoms. Less accuracy will be attained if only one symptom is entered. More the number of symptoms, the greater is the accuracy.

V. REFERENCES

- [1] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, "Prediction of heart disease using machine learning," in 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 1275–1278.
- [2] Y. Hasija, N. Garg, and S. Sourav, "Automated detection of dermatological disorders through image-processing and machine learning," in 2017, International Conference on Intelligent Sustainable Systems (ICISS), 2017, pp. 1047–1051.
- [3] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," BMC Medical Informatics and Decision Making, vol. 19, no. 1, pp. 1–16, 2019.
- [4] R. Katarya and P. Srinivas, "Predicting heart disease at early stages using machine learning: A survey," in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 302–305.
- [5] S. Jadhav, R. Kasar, N. Lade, M. Patil, and S. Kolte, "Disease Prediction by Machine Learning from Healthcare Communities," International Journal of Scientific Research in Science and Technology, pp. 29–35, 2019.