Risk Assess - risk assessment of diseases

Shreyas Ambre[1], Amol Gosavi[2], Shubhansh Jha[3], Shubham Kumar[4]

[1,2,3,4]*Student, Computer Engineering, MCT's Rajiv Gandhi Institute of Technology, Mumbai University*

*Abstract-***Disease Prediction system is based on predictive modelling predicts the disease of the user on the basis of the symptoms that user provides as an input to the system. The system analyzes symptoms provided by the user as input and gives the probability of the disease as an output Disease Prediction is done. We are using different machine learning algorithm because it shows significant potential in surpassing standard system for disease diagnosis and medical expect in early detection of high risk-disease. This model predicts life threatening disease like cancer, HIV and Diabetes etc.**

I. INTRODUCTION

Machine learning is programming computers to optimize a performance using example data or past data. Infectious diseases have risen recently within increase in population. This model predicts the output on the basis of symptoms that the user will enter in the system and more accurate prediction will by the system. To detect the various disease, this system uses different machine learning algorithm like random forest, decision tree and naïve Bayes' algorithm. In this 21st Century humans are surrounded with technology as they avoid to go to hospital for small problem which may become a major disease in future. Multiple Disease prediction will give more accurate results, better performance and time efficient.

II. OBJECTIVE

There is need to study and make a system which is easy to operate by end user. This system also helps to predicts life threatening disease at early stage of life without visiting doctor for diagnosis. This system is time efficient, better performance and give more accurate disease predictions. User can ask queries related to their health that can answered by doctor. Our system also provides medicines facility and advice from expert doctors.

III. PROBLEM STATEMENT

Chronic disease is a major contributor to the burden of diseases. People tend to ignore the early symptoms of a disease as they are mild. And till the time they become aware of the disease, the infection has already spread into their body. In today environment due to day-to-day busy schedule many times it is not possible for the patients to go and get a medical appointment booked in person. For small problem, user have to go personally to the hospital for check-up which is more time consuming. People ignore such model because the system is more complex to operate and had to pay price.

IV. LITERATURE REVIEW

A. Existing System

Risk assessment is one of the most trusted and the popular names in the medical appointment making system using the medium of internet. This system not only predicts the disease but also provide various services like doctor appointment, solving patient health related question by expert doctors, medicine facility and free check-up. Our system will provide service 24/7.

B. Proposed system

We are proposing such a system which is simple and user friendly and time efficient. Our aim with this system is to build connection between doctors and patients. Patients can ask any queries related to their health that can answered immediately by expert. The main feature will be the machine learning, which will help us in getting accurate predictions. After predicting the disease, our system will suggest patients to get doctors consultation on the report.

IV. SYSTEM ARCHITECTURE
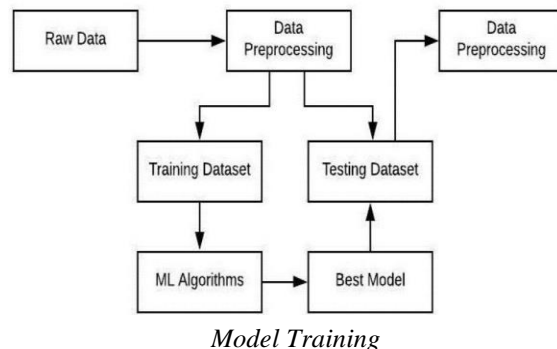

Model Training

V. DATASET AND MODEL DESCRIPTION

We have taken the sample dataset from Kaggle. In this dataset which is used to train the machine learning model. The dataset will contain symptoms of various diseases. In our machine learning model patient can input five different symptoms and get resulted disease based on three algorithms.
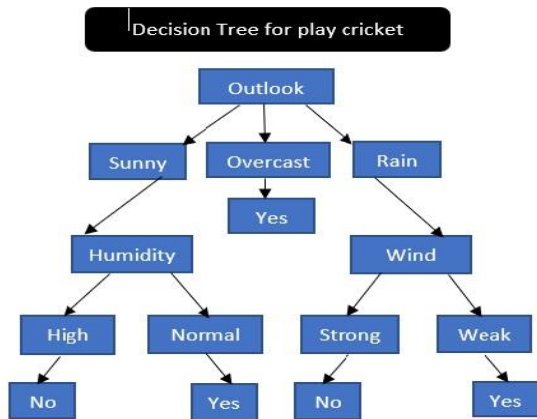
VI. ALGORITHMS

A. Random Forest

Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.
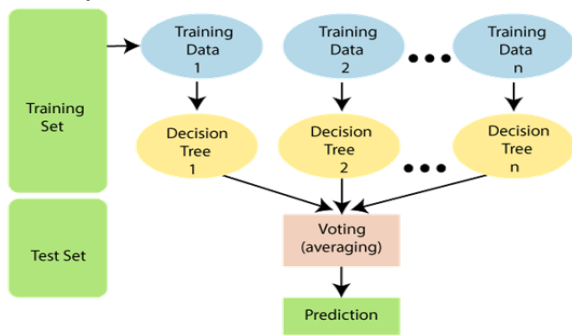
B. Decision Tree

Decision tree is flowchart-like tree structure, where internal node represents a test on an attribute, each branch represents outcome of the test and each terminal node is class label.


C. Naïve Bayes

Naïve bayes algorithm is supervised machine learning algorithm, used for solving classification problem. This algorithm is also known as probabilistic classifier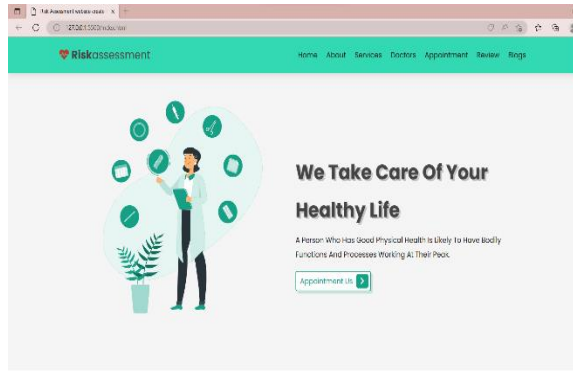 because it predicts output on the basis of the probability of object. Naïve means that the occurrence of certain feature is different or independent of occurrences of other features. It called Bayes' because it is based on principle of Bayes' theorem.

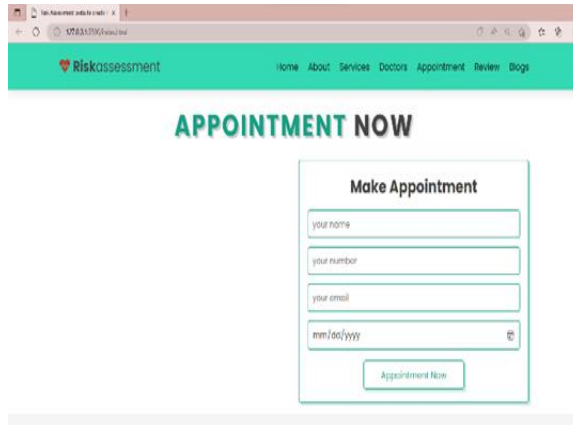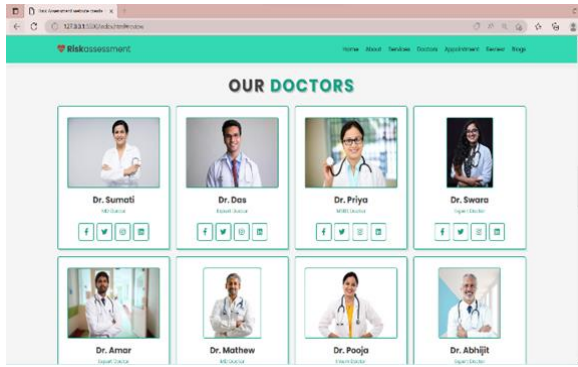$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

VII. RESULT
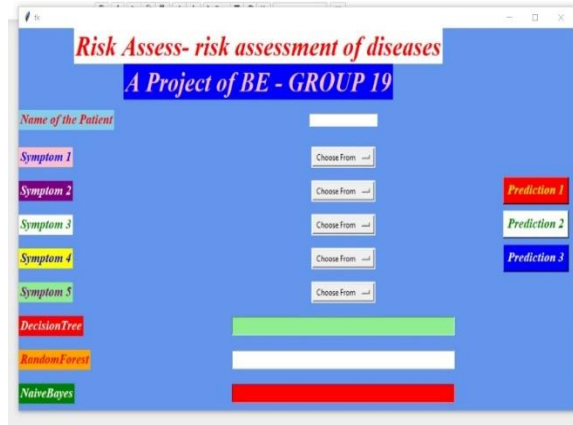
Disease Prediction

Home Page

Book Appointment

Doctors Available

VIII. METHODOLOGY

We are using Machine learning algorithm for disease predictions. We have taken sample dataset from Kaggle. Then apply Operation like data preprocessing, data cleaning on sample dataset for further training and testing data. After training and testing dataset will show prediction of disease as an output. The patient has to fill the input symptoms asked in system. The algorithm like random forest, decision tree and naïve bayes work simultaneously on given input symptoms and give separate predictive result in the form of disease to the concern patient. Patient can book appointment for consulting the doctors for their respective disease.

IX. CONCLUSION

This project aims to predict the disease on the basis of the symptoms. The project is designed in such a way that the system takes some symptoms from the user as input and produces output according to the algorithm use in project. Our system provide user friendly environment.

REFERENCE

[1] Min Chen, Yixue Hao, Kai Hwang, Fellow, IEEE, Lu Wang, and Lin Wang "Disease Prediction by Machine Learning over Big Data from Healthcare Communities" (2017).

[2] Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.

[3] P. Groves, B. Kayyali, D. Knott, and S. V. Kuiken, "The 'big data' revolution in healthcare: Accelerating value and innovation," 2016.

[4] S.-H. Wang, T.-M. Zhan, Y. Chen, Y. Zhang, M. Yang, H.-M. Lu, H.-N. Wang, B. Liu, and P. Phillips, "Multiple sclerosis detection based on biorthogonal wavelet transform, rbf kernel principal component analysis, and logistic regression," IEEE Access, vol. 4, pp. 7567–7576, 2016.

[5] S. Patel and H. Patel, "Survey of data mining techniques used in healthcare domain," Int. J. of Inform. Sci. and Tech., Vol. 6, pp. 53-60, March 2016.