A

Mini project Disease Prediction System

Using Machine Learning

Submitted By: Chirag Aggarwal B.Tech (CSE)

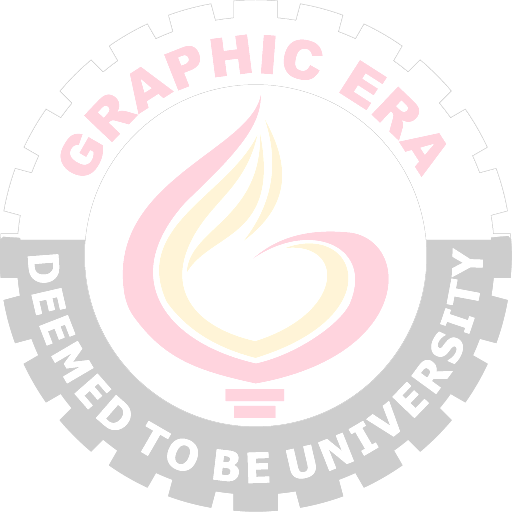
## 4TH SEM

Under The Guidance of

## DR. Manoj Diwaker

PROFESSOR Department of Computer Science

# CANDIDATE’S DECLARATION

I hereby declare that the mini project work being presented in this report

entitled “Disease Prediction System Using Machine Learning”

# submitted in the department of computer science, faculty of technology , Graphic Era University, Dehradun is the authentic work carried out by me under the guidance of DR. Manoj Diwaker, Department of computer science, Graphic Era University, Dehradun.

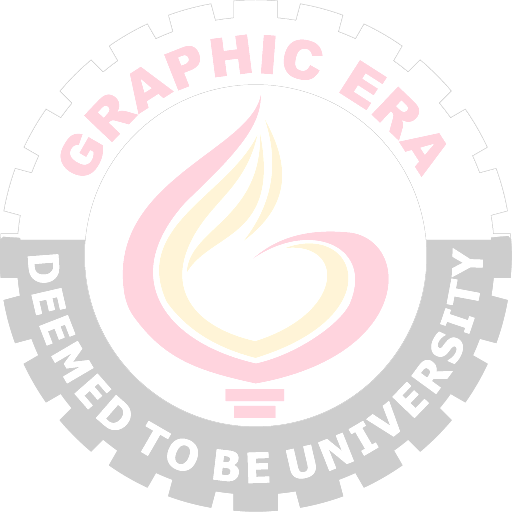
DATE : 14/July/2022 CHIRAG AGGARWAL

B.TECH (CSE)

4th SEM

Graphic Era University Dehradun

## Abstracts:

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 aim of developing classifier system using machine learning algorithms is to immensely help to solve the health related issues.

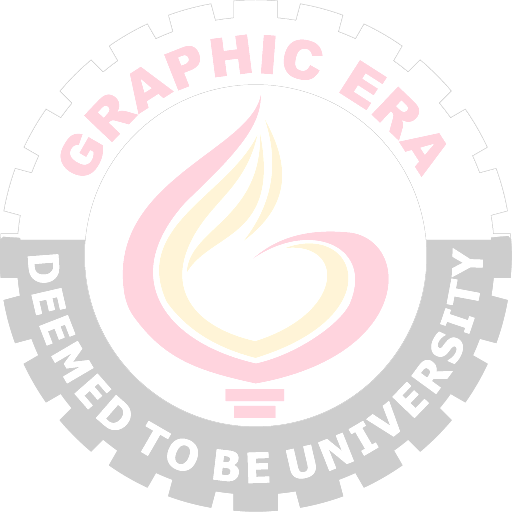
We collected a wider collection of data containing 816 drugs and 1393 diseases with their features .

This research work carried out demonstrates the disease prediction system developed using machine learning algorithms such as naïve bayes classifier..

Naïve Bayes classifier is used in the prediction of the disease which is a supervised machine learning algorithm. The probability of the disease is calculated by the Naïve Bayes algorithm. With an increase in biomedical and healthcare data, accurate analysis of medical data benefits early disease detection and patient care. By using linear regression and decision tree we are predicting diseases like Diabetes, Malaria, Jaundice, Dengue, and Tuberculosis.

Keywords: Disease Prediction, Machine learning, Naive bayes algorithm

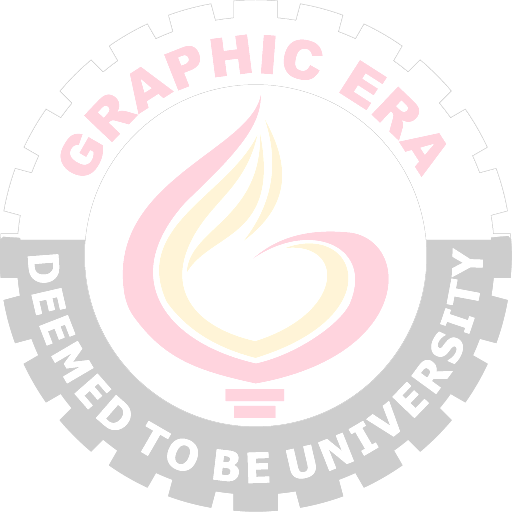
## INTRODUCTION

Machine Learning is the domain that uses past data for predicting. Machine Learning is the understanding of computer system under which the Machine Learning model learn from data and experience. The machine learning algorithm has two phases: 1) Training & 2) Testing. To predict the

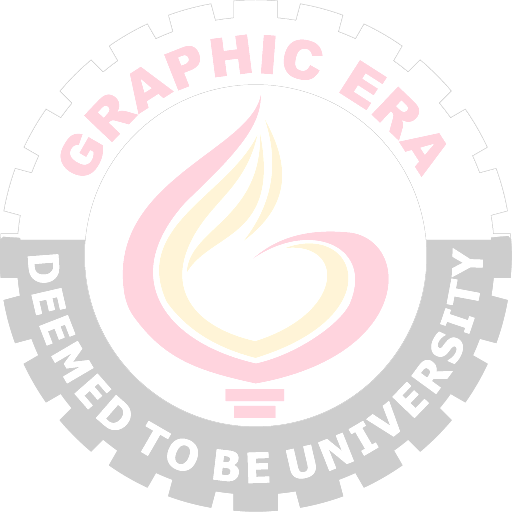
disease from a patient’s symptoms and from the history of the patient, machine learning technology is struggling from past decades. Healthcare issues can be solved efficiently by using Machine Learning Technology. We are applying complete machine

learning concepts to keep the track of patient’s health. ML model allows us to build models to get quickly cleaned and processed data and deliver results faster. By using this system doctors will make good decisions related to patient diagnoses and according to that, good treatment will be given to the patient, which increases improvement in patient healthcare services. To introduce machine learning in the medical field, healthcare is the prime example. To improve the accuracy of large data, the existing work will be done on unstructured or textual data

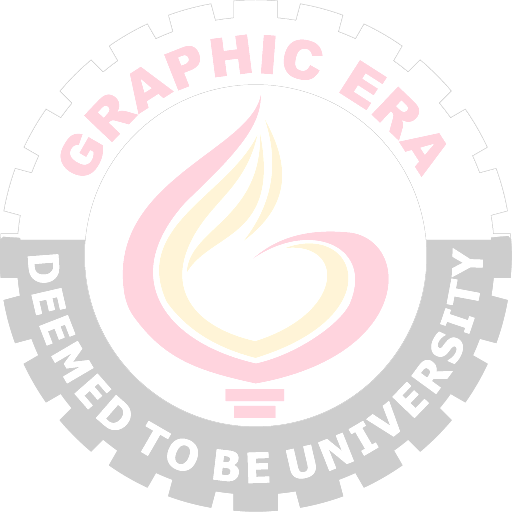
Motivation

The emergence of Artificial Intelligence (AI) enabled computerized systems to perceive, think and operate in an intelligent manner like humans. . ML algorithms apply various optimization, statistical ,and probabilistic techniques to learn from data that was generated from past experiences, and deploy it in decision making. These algorithms deemed to be applied in many disciplines including network intrusion recognition, customer purchase behavior detection, process manufacturing optimization, credit card fraud detection, and disease modulation. This presents the hypothesis that medical doctors can utilize supervised learning as a powerful tool to conduct diseases diagnosis more efficiently. Medicaid services and centers for Medicare reported that 50% of Americans had multiple chronic diseases, which led the US health care to spend around $3.3 trillion in 2016, that amounts to $10,348 per person in the US [5]. Moreover, the World Health Organization and World Economic Forum reported that India had a huge loss of $236.6 billion by 2015 because of fatal diseases, caused by malnutrition and morbid lifestyles . Such expenditures revealed how prone people are to a spectrum of diseases, which showcased how vital it is to detect diseases early, to consequently reduce the fatality of these maladies. In addition, early disease prediction can lessen the financial pressure on the economy and ensure better maintenance on the overall well-being of the community.

## AIM

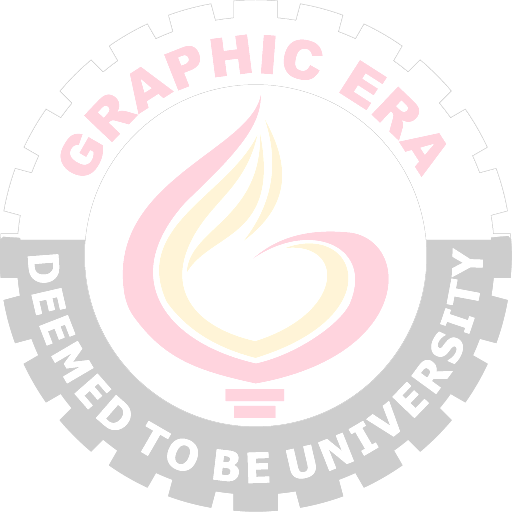
The aim of this study is to test the proposed hypothesis that supervised ML algorithms can improve health care by the accurate and early detection of diseases. In this study, we investigate studies that utilize more than one supervised ML model for each disease recognition problem. This approach renders more comprehensiveness and precision because the evaluation of the performance of a single algorithm over various study settings induces bias which generates imprecise results. The analysis of ML models will be conducted on few diseases located at heart, kidney, breast, and brain. For the detection of the disease.

METHODOLOGY

1. Data Collection And Data Analysis.
   * Training Data: Training data is also known as training datasets, training sets, and training sets. It is an important aspect of the machine learning model which helps us to make accurate predictions and perform the tasks we want. Simply put, training data forms a machine learning model and tells you what the awaited result looks like. The model iteratively analyzes the dataset to understand its attributes precisely and make appropriate changes to enhance the performance.
   * Testing Data: The test dataset is a subset of the training dataset used to make an objective evaluation of the final model.
2. Algorithms TECHNIQUES
   * NAIVE BAYES: Naive Bayes is an easy however amazingly powerful rule for prognosticative modeling. The independence assumption that allows decomposing joint likelihood into a product of marginal likelihoods is called as 'naive'. This simplified Bayesian classifier is called as naive Bayes. The Naive Bayes classifier assumes the presence of a particular feature in a class is unrelated to the presence of any other feature. It is very easy to build and useful for large datasets. Naive Bayes is a supervised learning model. Bayes theorem provides some way of calculative posterior chance P(b|a) from

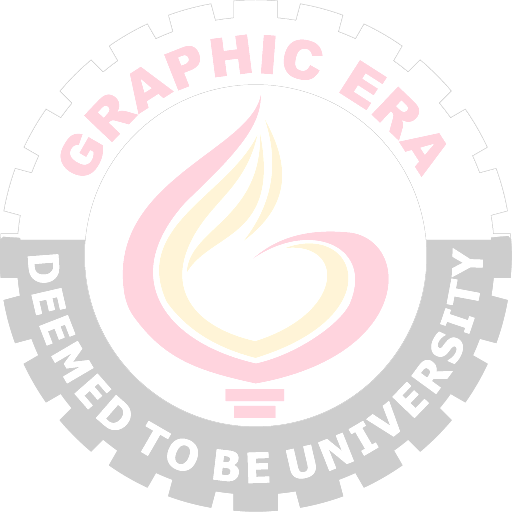
P(b), P(a) and P(a|b). Look atithe equation below: P(b v a)= P(a v b)P(b)/P(a)

Above,

* + P(b|a) is that the posterior chance of class (b,target) given predictor (a, attributes).
  + P(b) is the prioriprobability of class.
  + P(a|c)iis that chance that is that the chance of predictor given class.
  + P(a) is the prioriprobability of predictor. In our system, Naïve Bayes decides which symptom is to put in classifier and which is not. 8.3 LOGISTIC REGRESSION Logistic regression could be a supervised learning classification algorithm accustomed to predict the chance of a target variable that is Disease.

1. Tool Used :
2. Jupyter Notebook(Anaconda 3)
3. Gcc compiler
4. Language used
   * Python (ML)

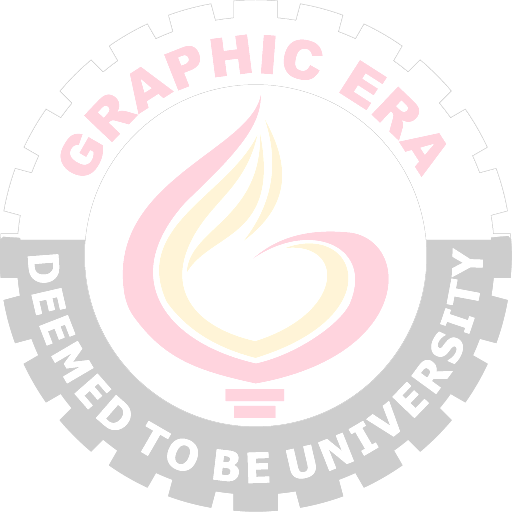
Modules

* + Main menu module
  + Display the result
  + Selects symptoms
  + Edit symptoms

### Features Of Payroll System:

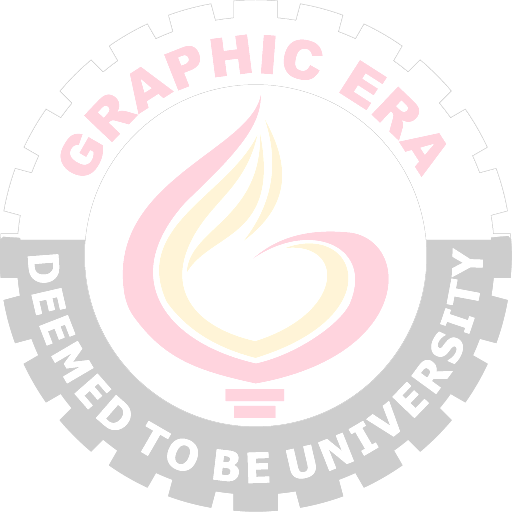
* + Easy to use.
  + Both Patients and doctors are use.
  + It is very interactive and save time.
  + It is completely secure.
  + Reduces the patients problems.

Conclusion :

The main aim of this disease prediction system is to predict the disease on the basis of the symptoms. This system takes the symptoms of the user from which he or she suffers as input and generates final output as a prediction of disease. Average prediction accuracy probability of 100% is obtained. Disease Predictor was successfully implemented using the grails framework. This system gives a user-friendly environment and easy to use.

As the system is based on the web application, the user can use this system from anywhere and at any time. In conclusion, for disease risk modeling, the accuracy of risk prediction depends on the diversity feature of the hospital data.This systematic review aims to determine the performance, limitations, and future use of Software in healthcare. Findings may help inform future developers of Disease Predictability Software and promote personalized patient care. The program predicts Patient Diseases. Disease Prediction is done through User Symbols.

### Bibliography:

* + [www.tutorialspoint.com](http://www.tutorialspoint.com/)
  + [www.youtube.com](http://www.youtube.com/)
  + [www.w3school.com](http://www.w3school.com/)