মেট্রোপলিটন ইউনিভার্সিটি
Metropolitan University

Department of Computer Science and Engineering

Project Report

Project Name: Multiple Disease Prediction System

Submitted By

Bipropod Das Shubro

Identity: 171-115-071

Batch: CSE 45th

Supervisor

Mr. Md Mushtaq Shahriyar Rafee

Lecturer

Metropolitan University, Sylhet

Date: November 22, 2022

ACKNOWLEDGEMENT

First of all, I want to thank the almighty God for allowing me to complete and submit the Final-year Project. I would not be able to submit the final year project without the blessings of Almighty God.

I would like to thank my Project Supervisor **Mr. Md Mushtaq Shahriyar Rafee,** Lecturer, Department of Computer Science and Engineering, Metropolitan University for guiding me to complete the Project. Without his supervision, this project could not be done. His involvement in this Project can only be acknowledged but never be rewarded.

Also, I would like to thank the Head of the Department **Mr**. **Md. Mahfujul Hasan** and all the faculty members of Department of Computer Science and Engineering, Metropolitan University for making me skillful and giving me an entrance in the field of Computer Science and Engineering.

Last but not the least, I would like to thank my Parents and other Teachers and some of my close friends for support throughout the Project.

CERTIFICATE

The Project titled **Multiple Disease Prediction System** is created by Bipropod Das Shubro, Id: 171-115-071, Batch: CSE-45th has been accepted for the partial fulfillment of the requirements for the degree of Bachelor of Computer Science and Engineering in August 2022.

Mr. Md Mushtaq Shahriyar Rafee

Project Supervisor

Lecturer

Department of Computer Science and Engineerin

Metropolitan University, Sylhet

ABSTRACT

In that Technology governed world of object-oriented development, the life-span of people in general has lengthened from the past-centuries, but the ratio of epidemic and various deadly diseases are taking a place of constant anxiety and uncertainty day by day. The range of disease are increasing hand in hand with the increase of modern technology and machinery enrichments. It is high time to make the Machine-Learning and Artificial Intelligence dominions humanistic and subjective as much as possible. By using these technological efficiencies the life-taking diseases are being detected and treated dynamically. Chronic diseases such as Heart disease, Cancer, Diabetes, and Parkinson's disease are the leading causes of disability and death in Bangladesh and throughout the world. In 2014, 8.5% of adults aged 18 years and older had diabetes. In 2019, diabetes was the direct cause of 1.5 million deaths and 48% of all deaths due to diabetes occurred before the age of 70 years. Another 460 000 kidney disease deaths were caused by diabetes, and raised blood glucose causes around 20% of cardiovascular deaths *(1)*. According to the latest WHO data published in 2020 Coronary Heart Disease Deaths in Bangladesh reached 108,528 or 15.16% of total deaths and Parkinson's disease Deaths reached 3,782 or 0.53% of total deaths also Breast Cancer Deaths reached 6,808 or 0.95% of total deaths. As compare to other diseases these types of diseases having high rate of deaths, so there is need of promising solution over chronic diseases. Medical data growth in healthcare communities, accurate analysis of medical data benefit early disease detection, patient care and community services. However, the analysis of patients is depending on accuracy of diagnosis and then treatment as well. The wrong diagnosed patients lead to deaths in chronic type diseases. So the high risk of diagnosis there is a need of accurate diagnosis aid for chronic diseases. So we are proposing diagnosis system based on machine learning for giving promising solution with high accuracy. The proposed system consists of many diseases such as Breast Cancer, Parkinson disease, diabetes, heart disease detections and stages predictions. High rate of deaths due to chronic diseases such as heart disease, diabetes, Parkinson's disease, breast cancer need to develop proper diagnosis system which helps to doctors.. Many works is already carried out for different diseases but there is not any promising solution found that gives accurate diagnosis for all in one. The proposed system consists of many diseases such as lung cancer, brain tumor, heart disease detections and stages predictions. We are trying to develop system for multi disease detection and stages predictions gives early detection and saves lots of life's by reducing death rate by chronic diseases.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURE

ABBREVIATION

SVM	=	Support Vector Machine
NB	=	Naïve bayes
XGBoost	=	extreme Gradient Boosting
SSL	=	Self-Supervised Learning
SLP	=	Single-Layer Perceptron
MLP	=	Multi-Layer Perceptron
RL	=	Reinforcement Learning
NN	=	Neural Network
ML	=	Machine Learning
k-NN	=	k-Nearest Neighbor
df	=	Data Frame
CART	=	Classification And Regression Tree
CNN	=	Convolutional Neural Network
DeepLIFT	=	Deep Learning Important FeaTures
DL	=	Deep Learning
GA	=	Genetic Algorithm

CHAPTER - 1: INTRODUCTION

1.1 About

The medical services industry can go with a successful choice by “Data mining" the huge data set they have for example by extracting the hidden relationships and connections in the data set. Data mining algorithms like Decision Tree, Random Forest and Naïve Bayes calculations can give a solution for this present circumstance. Thus, we have developed a computerized framework that can discover and extract hidden knowledge associated with the diseases from a historical (diseases-side effects) data set by the standard arrangement of the particular algorithm. The medical care and clinical area are more in need of data mining today. At the point when certain information mining strategies are utilized in a correct manner, significant data can be removed from enormous data sets and that can assist the clinical specialist with taking early choice and further develop healthcare administrations. The spirit is to use the classification in order to assist the physician. During a ton of examinations over existing frameworks in medical services, examination thought about just a single sickness at a time. Most extreme articles center on a specific sickness. At the point when any association needs to break down their patient's wellbeing reports then they need to send many models. The methodology in the current framework is helpful to dissect just specific illnesses. These days mortality has expanded because of not distinguishing the specific infection. Indeed, even the patient who got restored from one sickness might be experiencing another infection. Inside experiencing heart issues which are not distinguished. Like this many occasions are seen in many individuals' life stories. In numerous sickness expectation frameworks a client can break down more than one illness on a solitary site. The client doesn't have to cross better places to foresee whether he/she has a specific infection or not. In this, the client needs to input the value of the specific tests information and some basic data and simply click on submit. The comparing Machine Learning model will be summoned and it will anticipate the result and show it on the screen with accuracy percentage.

1.2 Data Mining and Machine Learning Algorithm

The Data Mining and the Machine Learning Algorithms are used for the prediction of Diseases in the Project. There are different Data Mining and Machine Learning Algorithms used for the purpose of correcting and evaluating the dataset and then testing the dataset on the basis of train score and the test score of the ML model.

1.2.1 Data Analysis and Data Mining

The Data Mining is a process in which raw data is prepared and structured from the unstructured data as to take meaningful information from the data which can be used in the project. Task of making data organized and reflective about data is to

way to get what this information does the data contains in it and what it does not have in it. There are so many different types of methods in which the people can make use of data analysis. It is simply very easy to use data during the analysis phase and get to some certain conclusions or some agendas. The analysis of data is a process of inspecting, cleaning, transforming, and modeling data with the objective of highlighting useful information, suggesting conclusions, and supporting decision making which are helpful to the user. Data analysis has multiple facets and approaches, encompassing diverse techniques under an array of names, in different business, science, and social science domains.

Data Mining is the discovery of unknown information found in databases, data mining functions has some different methods for clustering, classification, prediction, and associations. In the data mining important application is that of mining association rules, association rules was first introduced in 1993 and are used to identify relationships among a set of items in databases these different properties are not based on the properties of the data, but rather based on co- occurrence of the data items. The Data mining helps in giving new and different perspectives for data analysis the main role of data mining is to extract and discover new knowledge from data. In the past few years, different methods have been coined and developed about the capabilities of data collection and data generation, data collection tools have provided us with a huge amount of data, data mining processes have integrated techniques from multiple disciplines such as, statistics, machine learning, database technology, pattern recognition, neural networks, information retrieval and spatial data analysis. The data mining techniques have been used in many different fields such as, business management, science, engineering, banking, data management, administration, and many other applications.

1.2.2 Machine Learning Algorithm:

The ML is a small part of Artificial Intelligence (AI) which is used in the computation work and the analysis work in the AI. The ML algorithms are used to find different patterns and different structures in the dataset which is provided to the dataset, the ML algorithms are used to give a large computation capabilities to the system by which a large amount of data is given to the model for the purpose of training and testing the data, the ML algorithms are used in decision making process the model which is prepared by using the ML has a large amount of data in it which makes it a very good for the process of decision making. ML algorithms have very high computational power and are proven to be very helpful in today's world.

Different types of ML algorithms are organized into different ways, based on the desired outcome of the algorithm. Common algorithm types include:

1. Supervised learning — The supervised learning algorithm can apply what has been learned in the past to new data using labelled examples to predict the future events. Starting from analysis of a known training dataset. This algorithm is used to provide targets for any new values after sufficient amount of training of the

model.

2. Unsupervised learning — Unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. This algorithm shows how the system can infer a function to describe a hidden structure from unlabeled data.

3. Semi-supervised learning — This category of the ML algorithms falls somewhere be- tween the supervised learning and the unsupervised learning algorithm which combines both labeled and unlabeled examples to generate an appropriate function or classifier which is used to make a model for the purpose of prediction or classification.

4. Reinforcement learning — This is the algorithm where the algorithm learns a policy of how to act given an observation of the world. Every action has some impact in the environment, and the environment provides feedback that guides the learning algorithm.

5. Transduction — This algorithm is similar to supervised learning, but does not explicitly construct a function: instead, tries to predict new outputs based on training inputs, training outputs, and new inputs.

6. Learning to learn — This method is where the algorithm learns its own inductive bias based on previous experience.

The performance and computational analysis of ML algorithms is a branch of statistics known as computational learning theory.

Machine learning is about designing algorithms that allow a computer to learn. Learning is not necessarily involves consciousness but learning is a matter of finding statistical regularities or other patterns in the data. Thus, many machine learning algorithms will barely resemble how human might approach a learning task. However, learning algorithms can give insight into the relative difficulty of learning in different environments

Machine learning is made up of three parts:

1. The computational algorithm at the core of making determinations.
2. Variables and features that make up the decision.
3. Base knowledge for which the answer is known that enables (trains) the system to learn.

Initially, the model is fed parameter data for which the answer is known. The algorithm is then run, and adjustments are made until the algorithm's output (learning) agrees with the known answer. At this point, increasing amounts of data are input to help the system learn and process higher computational decisions.

1.3Stream-lit Framework

Stream-lit is an open source app framework in Python language. It helps us create web apps for data science and machine learning in a short time. It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc. With Stream-lit, no callbacks are needed since widgets are treated as variables. Data caching simplifies and speeds up computation pipelines. Stream-lit watches for changes on updates of the linked Git repository and the application will be deployed automatically in the shared link.

1.3.1 Advantages of Stream-lit

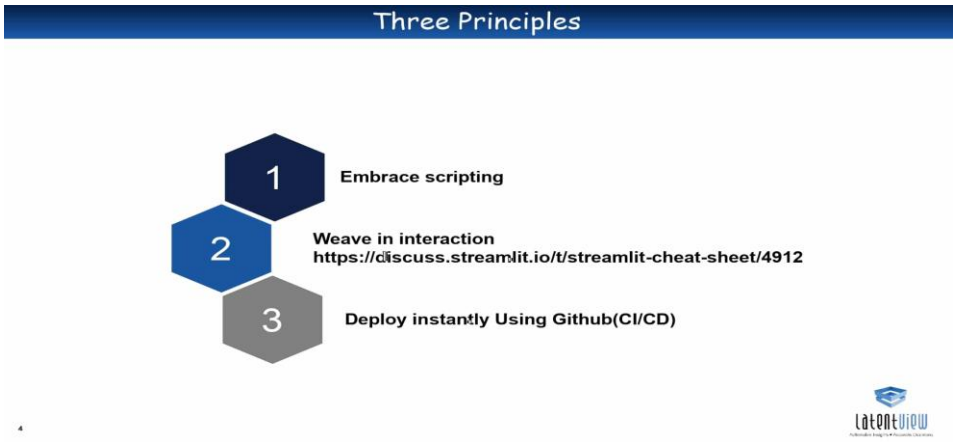

Fig-1: Principle of Stream-lit

Create a web app using python

- No need to write a separate backend. but we can utilize the other Server functionalities for data retrievals, defining the routes, handle HTTP requests, connect a frontend, write HTML, CSS, JavaScript,
- In Python script import the streamlit Like this
  - import streamlit as st
- Under imported **st** keyword we can create the line charts, map, table, baar charts, radio button, date picker, drag and drop options and etc…
  - st.line_chart("values")
- Using the below comment for running the script, The comment create the local web page.
  - **streamlit run scriptname.py**

5

1.3.2 Stream-lit Working

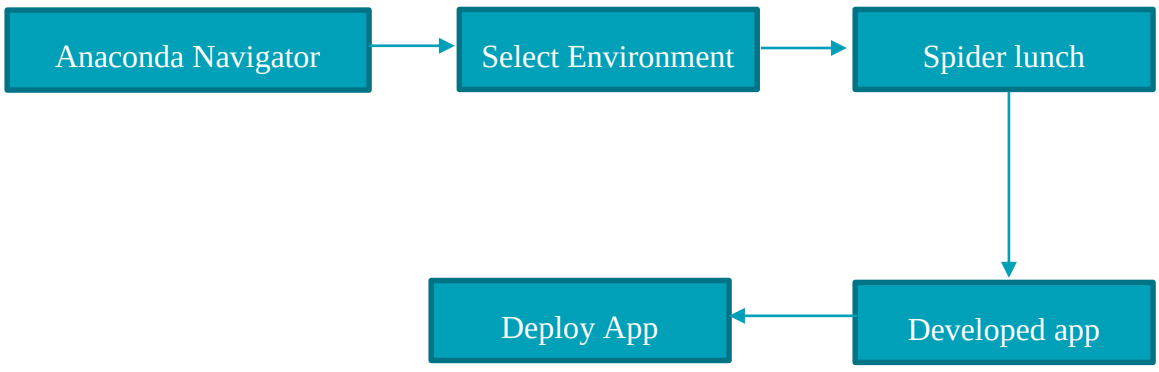

Fig-2: Stream-lit Working

Chapter 2: LITERATURE SURVEY

2.1 Background Study :

For diabetes detection patient do their pathology test, A1C Test, Fasting Blood Sugar Test, Glucose Tolerance Test, Random Blood Sugar Test, Glucose Screening Test, Glucose Tolerance Test.

Many different tests are used to diagnose heart disease. Besides blood tests and a chest X-ray, tests to diagnose heart disease can include: Electrocardiogram (ECG or EKG), Holter monitoring, Echocardiogram, Exercise tests or stress tests, Cardiac catheterization, Heart (cardiac) CT scan, Heart (cardiac) magnetic resonance imaging (MRI) scan.

Parkinson's disease can't be cured, but medications can help control the symptoms, often dramatically. In some more advanced cases, surgery may be advised. **Medication is the primal method to cure. The following methods of Medication are:**

Carbidopa-levodopa. (Rytary, Sinemet, Duopa, others), Levodopa, the most effective Parkinson's disease medication, is a natural chemical that passes into your brain and is converted to dopamine.Levodopa is combined with carbidopa (Lodosyn), which protects levodopa from early conversion to dopamine outside your brain. This prevents or lessens side effects such as nausea.

Inhaled carbidopa-levodopa. Inbrija is a brand-name drug delivering carbidopa-levodopa in an inhaled form. It may be helpful in managing symptoms that arise when oral medications suddenly stop working during the day.

Carbidopa-levodopa infusion. Duopa is a brand-name medication combining carbidopa and levodopa. However, it's administered through a feeding tube that delivers the medication in a gel form directly to the small intestine.

Catechol O-methyltransferase (COMT) inhibitors. Entacapone (Comtan) and opicapone (Ongentys) are the primary medications from this class. This medication mildly prolongs the effect of levodopa therapy by blocking an enzyme that breaks down dopamine.

Side effects, including an increased risk of involuntary movements (dyskinesia), mainly result from an enhanced levodopa effect. Other side effects include diarrhea, nausea or vomiting.

Deep brain stimulation. In deep brain stimulation (DBS), surgeons implant electrodes into a specific part of the brain. The electrodes are connected to a generator implanted in your chest near your collarbone that sends electrical pulses to your brain and may reduce your Parkinson's disease symptoms.

Tests and procedures used to diagnose breast cancer include:

- **Breast exam.** Your doctor will check both of your breasts and lymph nodes in your armpit, feeling for any lumps or other abnormalities.
- **Mammogram.** A mammogram is an X-ray of the breast. Mammograms are commonly used to screen for breast cancer. If an abnormality is detected on a screening mammogram, your doctor may recommend a diagnostic mammogram to further evaluate that abnormality.
- **Breast ultrasound.** Ultrasound uses sound waves to produce images of structures deep within the body. Ultrasound may be used to determine whether a new breast lump is a solid mass or a fluid-filled cyst.
- **Removing a sample of breast cells for testing (biopsy).** A biopsy is the only definitive way to make a diagnosis of breast cancer. During a biopsy, your doctor uses a specialized needle device guided by X-ray or another imaging test to extract a core of tissue from the suspicious area. Often, a small metal marker is left at the site within your breast so the area can be easily identified on future imaging tests.
- **Breast magnetic resonance imaging (MRI).** An MRI machine uses a magnet and radio waves to create pictures of the interior of your breast. Before a breast MRI, you receive an injection of dye. Unlike other types of imaging tests, an MRI doesn't use radiation to create the images.

Other tests and procedures may be used depending on your situation.

2.2 Paper Study:

Serial No	Paper Title	Author	Publication Year	Work Citation
1.	**Breast Cancer Detection Using Machine Learning Algorithms**	Shubham Sharma; Archit Aggarwal; Tanupriya Choudhury	2018	S. Sharma, A. Aggarwal and T. Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 2018, pp. 114-118, doi: 10.1109/CTEMS.2018.8769187.
2.	**Diabetes Detection Using Machine Learning Classification Methods**	Nour Abdulhadi; Amjed Al-Mousa	**2021**	N. Abdulhadi and A. Al-Mousa, "Diabetes Detection Using Machine Learning Classification Methods," 2021 International Conference on Information Technology (ICIT), 2021, pp. 350-354, doi: 10.1109/ICIT52682.2021.9491788.
3.	**Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease-A review**	Minja Belić , Vladislava Bobić , Milica Badža , Nikola Šolaja , Milica Đurić-Jovičić , Vladimir S Kostić	**2019**	Belić M, Bobić V, Badža M, Šolaja N, Đurić-Jovičić M, Kostić VS. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease-A review. Clin Neurol Neurosurg. 2019 Sep;184:105442. doi: 10.1016/j.clineuro.2019.105442. Epub 2019 Jul 16. PMID: 31351213.
4	**Heart Disease Prediction using Machine Learning Techniques**	Devansh Shah, Samir Patel & Santosh Kumar Bharti	**2020**	Shah, D., Patel, S. & Bharti, S.K. Heart Disease Prediction using Machine Learning Techniques. *SN COMPUT. SCI.* 1, 345 (2020). https://doi.org/10.1007/s42979-020-00365-y

Table-1: Paper Study

CHAPTER: 3: SYSTEM ANALYSIS

3.1 Major Inputs Required

The inputs required for the project are:

- Software Inputs:
  - Jupyter Notebook
  - Python version 3
  - Pip version 3
  - Pip virtual environment
  - Stream-lit version
  - Spider
  - Git hub
- Hardware Inputs:
  - Windows/Linux/Mac OS
  - At least of 2 GB RAM
  - At least 512 GB ROM
  - At least a Integrated Graphic card
- User Inputs:
  - Basic Details
  - Symptoms
  - Pathology results

3.2: Feasibility Analysis

3.2.1 Technical Feasibility

The project is technically feasible as it can be built using the existing

available technologies. Itis a web based applications that uses Stream-lit Framework. The technology required by Multiple Disease Prediction System is available and hence it is technically feasible.

3.2.2 Economic Feasibility

The project is economically feasible as the cost of the project is involved only in the hosting of the project. As the data samples increases, which consume more time and processing power. In that case better processor might be needed.

3.2.3 Operational Feasibility

The project is operationally feasible as the user having basic knowledge about computer and Internet. Multiple Disease Prediction is based on client-server architecture where client is users and serveris the machine where dataset and project are stored.

3.3 Algorithm Used:

In this project different ML algorithms are used and several techniques of data mining are alsoused to check the dataset that weather it is a balanced dataset or not and check for the data is structured or not for the disease prediction. The various ML algorithms used in this project are:

1. Logistic Regression
2. Decision Tree
3. Random forest
4. KNN
5. XGBoost

6. SVM
7. Gaussian Naïve Bayas

Introduction to Logistic Regression

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).

Mathematically, a logistic regression model predicts P(Y=1) as a function of X. It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

Introduction to SVM

Support vector machines (SVMs) are powerful yet flexible supervised machine learning algorithms which are used both for classification and regression. But generally, they are used in classification problems. In 1960s, SVMs were first introduced but later they got refined in 1990. SVMs have their unique way of implementation as compared to other machine learning algorithms. Lately, they are extremely popular because of their ability to handle multiple continuous and categorical variables.

Introduction to Decision Tree

In general, Decision tree analysis is a predictive modelling tool that can be applied across many areas. Decision trees can be constructed by an algorithmic approach that can split the dataset in different ways based on different conditions. Decisions tress are the most powerful algorithms that falls under the category of supervised algorithms.

They can be used for both classification and regression tasks. The two main entities of a tree are decision nodes, where the data is split and leaves, where we got outcome. The example of a binary tree for predicting whether a person is fit or unfit providing various information like age, eating habits and exercise habits, is given below −

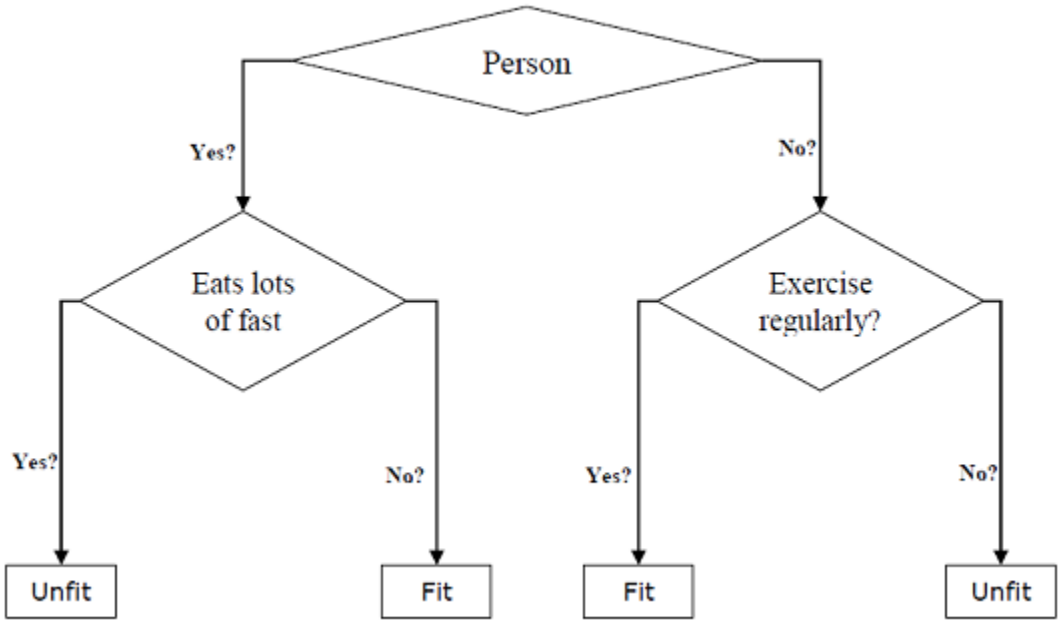

Fig-3: Decision Tree

In the above decision tree, the question are decision nodes and final outcomes are leaves. We have the following two types of decision trees −

- **Classification decision trees** − In this kind of decision trees, the decision variable is categorical. The above decision tree is an example of classification decision tree.
- **Regression decision trees** − In this kind of decision trees, the decision variable is continuous.

Introduction to Naïve Bayes Algorithm

Naïve Bayes algorithms is a classification technique based on applying Bayes' theorem with a strong assumption that all the predictors are independent to each other. In simple words, the assumption is that the presence of a feature in a class is independent to the presence of any other feature in the same class. For example, a phone may be considered as smart if it is having touch screen, internet facility, good camera etc. Though all these features are dependent on each other, they contribute independently to the probability of that the phone is a smart phone.

In Bayesian classification, the main interest is to find the posterior probabilities i.e. the probability of a label given some observed features, $(L \mid features)$. With the help of Bayes theorem, we can express this in quantitative form as follows −

P(L|features)=P(L)P(features|L)P(features)P(L|features)=P(L)P(features|L)$P(features)$

Here, ($L \mid features$) is the posterior probability of class.

(L) is the prior probability of class.

($features \mid L$) is the likelihood which is the probability of predictor given class.

($features$) is the prior probability of predictor.

Random forest introduction

Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result.

Introduction to k-means algorithm

K-means clustering algorithm computes the centroids and iterates until we it finds optimal centroid. It assumes that the number of clusters are already known. It is also called **flat clustering** algorithm. The number of clusters identified from data by algorithm is represented by 'K' in K-means.

In this algorithm, the data points are assigned to a cluster in such a manner that the sum of the squared distance between the data points and centroid would be minimum. It is to be understood that less variation within the clusters will lead to more similar data points within same cluster.

Introduction to Extreme Gradient Boosting (XGBoost)

When it comes to a superfast machine learning algorithm that works on tree-based models and tries to reach the best in class accuracy by optimally using computational resources, XGBoost or Extreme Gradient Boosting becomes the most natural choice. Created by Tianqi Chen, the XGBoost algorithm has recently got so much popularity owing to its massive usage in most of the hackathons and Kaggle competitions.

In simple terms, XGBoost may be formally defined as a decision tree-based ensemble learning framework that uses Gradient Descent as the underlying objective function and comes with a lot of flexibility while delivering the desired results by optimally using computational power.

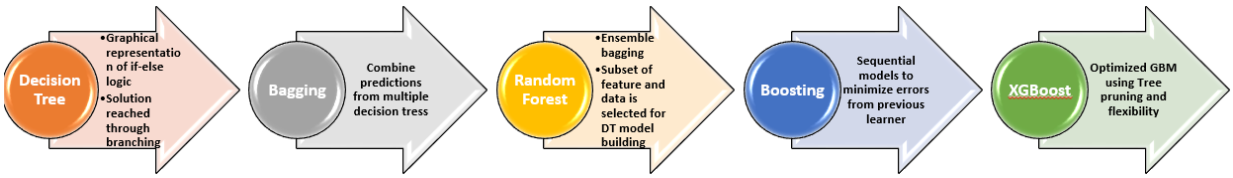

Figure 4 : The evolution of XGBoost from Tree-based models

Introduction To CNN (multi layered perceptron)

In **deep learning,** a **convolutional neural network** (**CNN/ConvNet**) is a class of **deep neural networks**, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics **convolution** is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

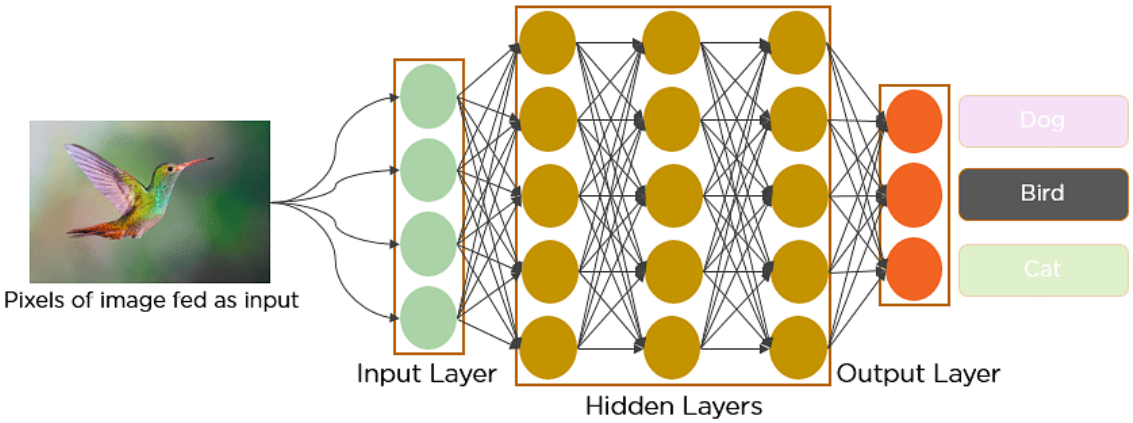

Fig-5: Multi layered Perceptron

But we don't really need to go behind the mathematics part to understand what a CNN is or how it works.

Bottom line is that the role of the ConvNet is to reduce the images into a form that is easier to process, without losing features that are critical for getting a good prediction.

CHAPTER 4: ABOUT DISEASES

4.1 Diabetes:

Diabetes

Key facts

- **The number of people with diabetes rose from 108 million in 1980 to 422 million in 2014. Prevalence has been rising more rapidly in low- and middle-income countries than in high-income countries.**
- **Diabetes is a major cause of blindness, kidney failure, heart attacks, stroke and lower limb amputation.**
- **Between 2000 and 2019, there was a 3% increase in diabetes mortality rates by age.**
- **In 2019, diabetes and kidney disease due to diabetes caused an estimated 2 million deaths.**
- **A healthy diet, regular physical activity, maintaining a normal body weight and avoiding tobacco use are ways to prevent or delay the onset of type 2 diabetes.**
- **Diabetes can be treated and its consequences avoided or delayed with diet, physical activity, medication and regular screening and treatment for complications.**

Overview

Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. Insulin is a hormone that regulates blood glucose. Hyperglycaemia, also called raised blood glucose or raised blood sugar, is a common effect of uncontrolled diabetes and over time leads to serious damage to many of the body's systems, especially the nerves and blood vessels.

In 2014, 8.5% of adults aged 18 years and older had diabetes. In 2019, diabetes was the direct cause of 1.5 million deaths and 48% of all deaths due to diabetes occurred before the age of 70 years. Another 460 000 kidney disease deaths were caused by diabetes, and raised blood glucose causes around 20% of cardiovascular deaths *(1)*.

Between 2000 and 2019, there was a 3% increase in age-standardized mortality rates from diabetes. In lower-middle-income countries, the mortality rate due to diabetes increased 13%.

By contrast, the probability of dying from any one of the four main noncommunicable diseases (cardiovascular diseases, cancer, chronic respiratory diseases or diabetes) between the ages of 30 and 70 decreased by 22% globally between 2000 and 2019.

Type 2 diabetes

Type 2 diabetes (formerly called non-insulin-dependent, or adult-onset) results from the body's ineffective use of insulin. More than 95% of people with diabetes have type 2 diabetes. This type of diabetes is largely the result of excess body weight and physical inactivity.

Symptoms may be similar to those of type 1 diabetes but are often less marked. As a result, the disease may be diagnosed several years after onset, after complications have already arisen.

Until recently, this type of diabetes was seen only in adults but it is now also occurring increasingly frequently in children.

Type 1 diabetes

Type 1 diabetes (previously known as insulin-dependent, juvenile or childhood-onset) is characterized by deficient insulin production and requires daily administration of insulin. In 2017 there were 9 million people with type 1 diabetes; the majority of them live in high-income countries. Neither its cause nor the means to prevent it are known.

Symptoms include excessive excretion of urine (polyuria), thirst (polydipsia), constant hunger, weight loss, vision changes, and fatigue. These symptoms may occur suddenly.

Gestational diabetes

Gestational diabetes is hyperglycemia with blood glucose values above normal but below those diagnostic of diabetes. Gestational diabetes occurs during pregnancy

Women with gestational diabetes are at an increased risk of complications during pregnancy and at delivery. These women and possibly their children are also at increased risk of type 2 diabetes in the future.

Gestational diabetes is diagnosed through prenatal screening, rather than through reported symptoms.

Impaired glucose tolerance and impaired fasting glycaemia

Impaired glucose tolerance (IGT) and impaired fasting glycaemia (IFG) are intermediate conditions in the transition between normality and diabetes. People with IGT or IFG are at high risk of progressing to type 2 diabetes, although this is not inevitable.

Health impact

Over time, diabetes can damage the heart, blood vessels, eyes, kidneys, and nerves.

- Adults with diabetes have a two- to three-fold increased risk of heart attacks and strokes *(2)*.
- Combined with reduced blood flow, neuropathy (nerve damage) in the feet increases the chance of foot ulcers, infection and eventual need for limb amputation.
- Diabetic retinopathy is an important cause of blindness and occurs as a result of long-term accumulated damage to the small blood vessels in the retina. Close to 1 million people are blind due to diabetes *(3)*.
- Diabetes is among the leading causes of kidney failure *(4)*.
- People with diabetes are more likely to have poor outcomes for several infectious diseases, including COVID-19.

Prevention

Lifestyle measures have been shown to be effective in preventing or delaying the onset of type 2 diabetes. To help prevent type 2 diabetes and its complications, people should:

- achieve and maintain a healthy body weight;
- be physically active – doing at least 30 minutes of regular, moderate-intensity activity on most days. More activity is required for weight control;
- Eat a healthy diet, avoiding sugar and saturated fats; and
- Avoid tobacco use – smoking increases the risk of diabetes and cardiovascular disease.

Diagnosis and treatment

Early diagnosis can be accomplished through relatively inexpensive testing of blood glucose.

Treatment of diabetes involves diet and physical activity along with lowering of blood glucose and the levels of other known risk factors that damage blood vessels. Tobacco use cessation is also important to avoid complications.

Interventions that are both cost-saving and feasible in low- and middle-income countries include:

- blood glucose control, particularly in type 1 diabetes. People with type 1 diabetes require insulin, people with type 2 diabetes can be treated with oral medication, but may also require insulin;
- blood pressure control; and
- foot care (patient self-care by maintaining foot hygiene; wearing appropriate footwear; seeking professional care for ulcer management; and regular examination of feet by health professionals).

Other cost saving interventions include:

- screening and treatment for retinopathy (which causes blindness);
- blood lipid control (to regulate cholesterol levels);
- screening for early signs of diabetes-related kidney disease and treatment.

4.2 Heart Disease

Overview

- Heart disease describes a range of conditions that affect the heart. Heart diseases include:
- Blood vessel disease, such as coronary artery disease
- Irregular heartbeats (arrhythmias)
- Heart problems you're born with (congenital heart defects)
- Disease of the heart muscle

- Heart valve disease
- Many forms of heart disease can be prevented or treated with healthy lifestyle choices.

Symptoms

Heart disease symptoms depend on the type of heart disease.

Symptoms of heart disease in the blood vessels

- Coronary artery disease is a common heart condition that affects the major blood vessels that supply the heart muscle. Cholesterol deposits (plaques) in the heart arteries are usually the cause of coronary artery disease. The buildup of these plaques is called atherosclerosis (ath-ur-o-skluh-ROE-sis). Atherosclerosis reduces blood flow to the heart and other parts of the body. It can lead to a heart attack, chest pain (angina) or stroke.
- Coronary artery disease symptoms may be different for men and women. For instance, men are more likely to have chest pain. Women are more likely to have other symptoms along with chest discomfort, such as shortness of breath, nausea and extreme fatigue.
- Symptoms of coronary artery disease can include:
- Chest pain, chest tightness, chest pressure and chest discomfort (angina)
- Shortness of breath
- Pain in the neck, jaw, throat, upper belly area or back
- Pain, numbness, weakness or coldness in the legs or arms if the blood vessels in those body areas are narrowed
- You might not be diagnosed with coronary artery disease until you have a heart attack, angina, stroke or heart failure. It's important to watch for heart symptoms and discuss concerns with your health care provider. Heart (cardiovascular) disease can sometimes be found early with regular health checkups.

Heart disease symptoms caused by irregular heartbeats (heart arrhythmias)

- The heart may beat too quickly, too slowly or irregularly. Heart arrhythmia symptoms can include:
- Chest pain or discomfort
- Dizziness
- Fainting (syncope) or near fainting
- Fluttering in the chest
- Lightheadedness
- Racing heartbeat (tachycardia)
- Shortness of breath
- Slow heartbeat (bradycardia)

Heart disease symptoms caused by congenital heart defects

- Serious congenital heart defects usually are noticed soon after birth. Congenital heart defect symptoms in children could include:
- Pale gray or blue skin or lips (cyanosis)
- Swelling in the legs, belly area or areas around the eyes
- In an infant, shortness of breath during feedings, leading to poor weight gain
- Less-serious congenital heart defects are often not diagnosed until later in childhood or during adulthood. Symptoms of congenital heart defects that usually aren't immediately life-threatening include:
- Easily getting short of breath during exercise or activity
- Easily tiring during exercise or activity
- Swelling of the hands, ankles or feet

Heart disease symptoms caused by diseased heart muscle (cardiomyopathy)

- Early stages of cardiomyopathy may not cause noticeable symptoms. As

the condition worsens, symptoms may include:

- Dizziness, lightheadedness and fainting
- Fatigue
- Feeling short of breath during activity or at rest
- Feeling short of breath at night when trying to sleep or waking up short of breath
- Irregular heartbeats that feel rapid, pounding or fluttering
- Swollen legs, ankles or feet

Heart disease symptoms caused by heart valve problems (valvular heart disease)

- The heart has four valves — the aortic, mitral, pulmonary and tricuspid valves. They open and close to move blood through the heart. Many things can damage the heart valves. A heart valve may become narrowed (stenosis), leaky (regurgitation or insufficiency) or close improperly (prolapse).
- Valvular heart disease is also called heart valve disease. Depending on which valve isn't working properly, heart valve disease symptoms generally include:
- Chest pain
- Fainting (syncope)
- Fatigue
- Irregular heartbeat
- Shortness of breath
- Swollen feet or ankles
- Endocarditis is an infection that affects the heart valves and inner lining of the heart chambers and heart valves (endocardium). Endocarditis symptoms can include:
- Dry or persistent cough
- Fever

- Heartbeat changes
- Shortness of breath
- Skin rashes or unusual spots
- Swelling of the legs or belly area
- Weakness or fatigue

When to see a doctor

- Seek emergency medical care if you have these heart disease symptoms:
- Chest pain
- Shortness of breath
- Fainting
- Always call 911 or emergency medical help if you think you might be having a heart attack.
- Heart disease is easier to treat when detected early. Talk to your health care provider if you have any concerns about your heart health. Together, you and your provider can discuss ways to reduce your heart disease risk. This is especially important if you have a family history of heart disease.
- If you think you may symptoms of heart disease, make an appointment to see your provider.

Causes

- **Heart disease causes depend on the specific type of heart disease. There are many different types of heart disease.**

How the heart works

- To understand the causes of heart disease, it may help to understand how the heart works.
- The heart is divided into chambers — two upper chambers (atria) and two lower chambers (ventricles).
- The right side of the heart moves blood to the lungs through blood vessels (pulmonary arteries).
- In the lungs, blood picks up oxygen and then returns to the left side of

the heart through the pulmonary veins.

- The left side of the heart then pumps the blood through the aorta and out to the rest of the body.

Heartbeats

- A beating heart squeezes (contracts) and relaxes in a continuous cycle.
- During contraction (systole), the lower heart chambers (ventricles) squeeze tight. This action forces blood to the lungs and the rest of the body.
- During relaxation (diastole), the ventricles fill with blood from the upper heart chambers (atria).

Electrical system

- The heart's electrical system keeps it beating. The heartbeat controls the continuous exchange of oxygen-rich blood with oxygen-poor blood. This exchange keeps you alive.
- Electrical signals start in the upper right chamber (right atrium).
- The signals travel through specialized pathways to the lower heart chambers (ventricles). This tells the heart to pump.

Causes of coronary artery disease

- A buildup of fatty plaques in the arteries (atherosclerosis) is the most common cause of coronary artery disease. Risk factors include a poor diet, lack of exercise, obesity and smoking. Healthy lifestyle choices can help lower the risk of atherosclerosis.

Causes of irregular heartbeats (arrhythmias)

- **Common causes of arrhythmias or conditions that can lead to them include:**
- Cardiomyopathy
- Coronary artery disease
- Diabetes
- Drug misuse

- Emotional stress
- Excessive use of alcohol or caffeine
- Heart problem present at birth (congenital heart defects)
- High blood pressure
- Smoking
- Heart valve disease
- Use of certain medications, including those bought without a prescription, and herbs and supplements

Causes of congenital heart defects

- A congenital heart defect develops while a baby is growing in the womb. A congenital heart defect forms as the baby's heart develops, about a month after conception. Congenital heart defects change the flow of blood in the heart. Some medical conditions, medications and genes increase the risk of congenital heart defects.

Causes of a thickened or enlarged heart muscle (cardiomyopathy)

- The cause of cardiomyopathy depends on the type:
- Dilated cardiomyopathy. The cause of this most common type of cardiomyopathy often is unknown. It may be passed down through families (inherited). Dilated cardiomyopathy typically starts in the heart's main pumping chamber (left ventricle). Many things can cause damage to the left ventricle, including heart attacks, infections, toxins and some drugs, including cancer medicines.
- Hypertrophic cardiomyopathy. This type is usually passed down through families (inherited).
- Restrictive cardiomyopathy. This is the least common type of cardiomyopathy. It can occur for no known reason. Sometimes it's caused by a buildup of protein called amyloid in the heart (cardiac amyloidosis) or connective tissue disorders.

Causes of heart valve disease

- Many things can cause diseases of the heart valves. Some people are born with heart valve disease (congenital heart valve disease). Heart valve disease may also be caused by conditions such as:
- Rheumatic fever
- Infections (infectious endocarditis)
- Connective tissue disorders

Risk factors

- Risk factors for heart disease include:
- Age. Growing older increases the risk of damaged and narrowed arteries and a weakened or thickened heart muscle.
- Sex. Men are generally at greater risk of heart disease. The risk for women increases after menopause.
- Family history. A family history of heart disease increases the risk of coronary artery disease, especially if a parent developed it at an early age (before age 55 for a male relative, such as your brother or father, and 65 for a female relative, such as your mother or sister).
- Smoking. If you smoke, quit. Substances in tobacco smoke damage the arteries. Heart attacks are more common in smokers than in nonsmokers. If you need help quitting, talk to your health care provider about strategies that can help.
- Unhealthy diet. Diets high in fat, salt, sugar and cholesterol have been linked to heart disease.
- High blood pressure. Uncontrolled high blood pressure can cause the arteries to become hard and thick. These changes interrupt blood flow to the heart and body.
- High cholesterol. Having high cholesterol increases the risk of atherosclerosis. Atherosclerosis has been linked to heart attacks and strokes.
- Diabetes. Diabetes increases the risk of heart disease. Obesity and high

blood pressure increase the risk of diabetes and heart disease.

- Obesity. Excess weight typically worsens other heart disease risk factors.
- Lack of exercises. Being inactive (sedentary lifestyle) is associated with many forms of heart disease and some of its risk factors, too.
- Stress. Unrelieved stress may damage the arteries and worsen other risk factors for heart disease.
- Poor dental health. It's important to brush and floss your teeth and gums often. Also get regular dental checkups. Unhealthy teeth and gums makes it easier for germs to enter the bloodstream and travel to the heart. This can cause endocarditis.

Complications

- **Complications of heart disease include:**
- **Heart failure.** This is one of the most common complications of heart disease. Heart failure occurs when the heart can't pump enough blood to meet the body's needs.
- **Heart attack.** A heart attack may occur if a blood clot is stuck in a blood vessel that goes to the heart.
- **Stroke.** The risk factors that lead to heart disease can also lead to an ischemic stroke. This type of stroke happens when the arteries to the brain are narrowed or blocked. Too little blood reaches the brain. A stroke is a medical emergency — brain tissue begins to die within just a few minutes of a stroke.
- **Aneurysm.** An aneurysm is a bulge in the wall of an artery. If an aneurysm bursts, you may have life-threatening internal bleeding.
- **Peripheral artery disease.** In this condition, the arms or legs — usually the legs — don't get enough blood. This causes symptoms, most notably leg pain when walking (claudication). Atherosclerosis can lead to peripheral artery disease.
- **Sudden cardiac arrest.** Sudden cardiac arrest is the sudden loss of heart

function, breathing and consciousness. It's usually due to a problem with the heart's electrical system. Sudden cardiac arrest is a medical emergency. If not treated immediately, it results in sudden cardiac death.

Prevention

- The same lifestyle changes used to manage heart disease may also help prevent it. Try these heart-healthy tips:
- Don't smoke.
- Eat a diet that's low in salt and saturated fat.
- Exercise at least 30 minutes a day on most days of the week.
- Maintain a healthy weight.
- Reduce and manage stress.
- Control high blood pressure, high cholesterol and diabetes.
- Get good sleep. Adults should aim for 7 to 9 hours daily.

4.3 Parkinson’s Disease

Parkinson's disease

Overview

- Parkinson's disease is a progressive disorder that affects the nervous system and the parts of the body controlled by the nerves. Symptoms start slowly. The first symptom may be a barely noticeable tremor in just one hand. Tremors are common, but the disorder may also cause stiffness or slowing of movement.

- In the early stages of Parkinson's disease, your face may show little or no expression. Your arms may not swing when you walk. Your speech may become soft or slurred. Parkinson's disease symptoms worsen as your condition progresses over time.

- Although Parkinson's disease can't be cured, medications might significantly improve your symptoms. Occasionally, your health care provider may suggest surgery to regulate certain regions of your brain and improve your symptoms.

Symptoms

- Parkinson's disease signs and symptoms can be different for everyone. Early signs may be mild and go unnoticed. Symptoms often begin on one side of the body and usually remain worse on that side, even after symptoms begin to affect the limbs on both sides.

- Parkinson's signs and symptoms may include:

- **Tremor.** A tremor, or rhythmic shaking, usually begins in a limb, often your hand or fingers. You may rub your thumb and forefinger back and forth. This is known as a pill-rolling tremor. Your hand may tremble when it's at rest. The shaking may decrease when you are performing tasks.
- **Slowed movement (bradykinesia).** Over time, Parkinson's disease may slow your movement, making simple tasks difficult and time-consuming. Your steps may become shorter when you walk. It may be difficult to get out of a chair. You may drag or shuffle your feet as you try to walk.
- **Rigid muscles.** Muscle stiffness may occur in any part of your body. The stiff muscles can be painful and limit your range of motion.
- **Impaired posture and balance.** Your posture may become stooped. Or you may fall or have balance problems as a result of Parkinson's disease.
- **Loss of automatic movements.** You may have a decreased ability to perform unconscious movements, including blinking, smiling or swinging your arms when you walk.
- **Speech changes.** You may speak softly, quickly, slur or hesitate before talking. Your speech may be more of a monotone rather than have the usual speech patterns.
- **Writing changes.** It may become hard to write, and your writing may appear small.

When to see a doctor

- See your health care provider if you have any of the symptoms associated with Parkinson's disease — not only to diagnose your condition but also to rule out other causes for your symptoms.

Causes

- In Parkinson's disease, certain nerve cells (neurons) in the brain gradually break down or die. Many of the symptoms are due to a loss of neurons that produce a chemical messenger in your brain called dopamine. When dopamine levels decrease, it causes atypical brain activity, leading to impaired movement and other symptoms of Parkinson's disease.

- The cause of Parkinson's disease is unknown, but several factors appear to play a role, including:

- **Genes.** Researchers have identified specific genetic changes that can cause Parkinson's disease. But these are uncommon except in rare cases with many family members affected by Parkinson's disease.

- However, certain gene variations appear to increase the risk of Parkinson's disease but with a relatively small risk of Parkinson's disease for each of these genetic markers.

- **Environmental triggers.** Exposure to certain toxins or environmental factors may increase the risk of later Parkinson's disease, but the risk is small.

- Researchers have also noted that many changes occur in the brains of people with Parkinson's disease, although it's not clear why these changes occur. These changes include:

- **The presence of Lewy bodies.** Clumps of specific substances within brain cells are microscopic markers of Parkinson's disease. These are called Lewy bodies, and researchers believe these Lewy bodies hold an important clue to the cause of Parkinson's disease.
- **Alpha-synuclein found within Lewy bodies.** Although many substances are found within Lewy bodies, scientists believe an important one is the natural and widespread protein called alpha-synuclein (a-synuclein). It's found in all Lewy bodies in a clumped form that cells can't break down. This is currently an important focus among Parkinson's disease researchers.

Risk factors

- Risk factors for Parkinson's disease include:

- **Age.** Young adults rarely experience Parkinson's disease. It ordinarily begins in middle or late life, and the risk increases with age. People usually develop the disease around age 60 or older. If a young person does have Parkinson's disease, genetic counseling might be helpful in making family planning decisions. Work, social situations and medication side effects are also different from those of an older person with Parkinson's disease and require special considerations.
- **Heredity.** Having a close relative with Parkinson's disease increases the chances that you'll develop the disease. However, your risks are still small unless you have many relatives in your family with Parkinson's disease.
- **Sex.** Men are more likely to develop Parkinson's disease than women.
- **Exposure to toxins.** Ongoing exposure to herbicides and pesticides may slightly increase your risk of Parkinson's disease.

Complications

- Parkinson's disease is often accompanied by these additional problems, which may be treatable:

- **Thinking difficulties.** You may experience cognitive problems (dementia) and thinking difficulties. These usually occur in the later stages of Parkinson's disease. Such cognitive problems aren't usually helped by medications.

- **Depression and emotional changes.** You may experience depression, sometimes in the very early stages. Receiving treatment for depression can make it easier to handle the other challenges of Parkinson's disease.

- You may also experience other emotional changes, such as fear, anxiety or loss of motivation. Health care providers may give you medication to treat these symptoms.

- **Swallowing problems.** You may develop difficulties with swallowing as your condition progresses. Saliva may accumulate in your mouth due to slowed swallowing, leading to drooling.
- **Chewing and eating problems.** Late-stage Parkinson's disease affects the muscles in the mouth, making chewing difficult. This can lead to choking and poor nutrition.
- **Sleep problems and sleep disorders.** People with Parkinson's disease often have sleep problems, including waking up frequently throughout the night, waking up early or falling asleep during the day.

- People may also experience rapid eye movement sleep behavior disorder, which involves acting out your dreams. Medications may improve your sleep.

- **Bladder problems.** Parkinson's disease may cause bladder problems, including being unable to control urine or having difficulty in urinating.
- **Constipation.** Many people with Parkinson's disease develop constipation, mainly due to a slower digestive tract.

- You may also experience:

- **Blood pressure changes.** You may feel dizzy or lightheaded when you stand due to a sudden drop in blood pressure (orthostatic hypotension).
- **Smell dysfunction.** You may experience problems with your sense of smell. You may have difficulty identifying certain odors or the difference between odors.
- **Fatigue.** Many people with Parkinson's disease lose energy and experience fatigue, especially later in the day. The cause isn't always known.
- **Pain.** Some people with Parkinson's disease experience pain, either in specific areas of their bodies or throughout their bodies.
- **Sexual dysfunction.** Some people with Parkinson's disease notice a decrease in sexual desire or performance.

Prevention

- Because the cause of Parkinson's is unknown, there are no proven ways to prevent the disease.

- Some research has shown that regular aerobic exercise might reduce the risk of Parkinson's disease.

- Some other research has shown that people who consume caffeine — which is found in coffee, tea and cola — get Parkinson's disease less often than those who don't

drink it. Green tea is also related to a reduced risk of developing Parkinson's disease. However, it is still not known whether caffeine protects against getting Parkinson's or is related in some other way. Currently there is not enough evidence to suggest that drinking caffeinated beverages protects against Parkinson's.

4.4 Breast Cancer

Breast cancer

Overview

Breast cancer is cancer that forms in the cells of the breasts.

After skin cancer, breast cancer is the most common cancer diagnosed in women in the United States. Breast cancer can occur in both men and women, but it's far more common in women.

Substantial support for breast cancer awareness and research funding has helped create advances in the diagnosis and treatment of breast cancer. Breast cancer survival rates have increased, and the number of deaths associated with this disease is steadily declining, largely due to factors such as earlier detection, a new personalized approach to treatment and a better understanding of the disease.

Symptoms

Signs and symptoms of breast cancer may include:

- A breast lump or thickening that feels different from the surrounding tissue
- Change in the size, shape or appearance of a breast
- Changes to the skin over the breast, such as dimpling
- A newly inverted nipple
- Peeling, scaling, crusting or flaking of the pigmented area of skin surrounding the nipple (areola) or breast skin
- Redness or pitting of the skin over your breast, like the skin of an orange

When to see a doctor

If you find a lump or other change in your breast — even if a recent mammogram was normal — make an appointment with your doctor for prompt evaluation.

Causes

Doctors know that breast cancer occurs when some breast cells begin to grow abnormally. These cells divide more rapidly than healthy cells do and continue to accumulate, forming

a lump or mass. Cells may spread (metastasize) through your breast to your lymph nodes or to other parts of your body.

Breast cancer most often begins with cells in the milk-producing ducts (invasive ductal carcinoma). Breast cancer may also begin in the glandular tissue called lobules (invasive lobular carcinoma) or in other cells or tissue within the breast.

Researchers have identified hormonal, lifestyle and environmental factors that may increase your risk of breast cancer. But it's not clear why some people who have no risk factors develop cancer, yet other people with risk factors never do. It's likely that breast cancer is caused by a complex interaction of your genetic makeup and your environment.

Inherited breast cancer

Doctors estimate that about 5 to 10 percent of breast cancers are linked to gene mutations passed through generations of a family.

A number of inherited mutated genes that can increase the likelihood of breast cancer have been identified. The most well-known are breast cancer gene 1 (BRCA1) and breast cancer gene 2 (BRCA2), both of which significantly increase the risk of both breast and ovarian cancer.

If you have a strong family history of breast cancer or other cancers, your doctor may recommend a blood test to help identify specific mutations in BRCA or other genes that are being passed through your family.

Consider asking your doctor for a referral to a genetic counselor, who can review your family health history. A genetic counselor can also discuss the benefits, risks and limitations of genetic testing to assist you with shared decision-making.

Risk factors

A breast cancer risk factor is anything that makes it more likely you'll get breast cancer. But having one or even several breast cancer risk factors doesn't necessarily mean you'll develop breast cancer. Many women who develop breast cancer have no known risk factors other than simply being women.

Factors that are associated with an increased risk of breast cancer include:

- **Being female.** Women are much more likely than men are to develop breast cancer.
- **Increasing age.** Your risk of breast cancer increases as you age.
- **A personal history of breast conditions.** If you've had a breast biopsy that found lobular carcinoma in situ (LCIS) or atypical hyperplasia of the breast, you have an increased risk of breast cancer.
- **A personal history of breast cancer.** If you've had breast cancer in one breast, you have an increased risk of developing cancer in the other breast.

- **A family history of breast cancer.** If your mother, sister or daughter was diagnosed with breast cancer, particularly at a young age, your risk of breast cancer is increased. Still, the majority of people diagnosed with breast cancer have no family history of the disease.
- **Inherited genes that increase cancer risk.** Certain gene mutations that increase the risk of breast cancer can be passed from parents to children. The most well-known gene mutations are referred to as BRCA1 and BRCA2. These genes can greatly increase your risk of breast cancer and other cancers, but they don't make cancer inevitable.
- **Radiation exposure.** If you received radiation treatments to your chest as a child or young adult, your risk of breast cancer is increased.
- **Obesity.** Being obese increases your risk of breast cancer.
- **Beginning your period at a younger age.** Beginning your period before age 12 increases your risk of breast cancer.
- **Beginning menopause at an older age.** If you began menopause at an older age, you're more likely to develop breast cancer.
- **Having your first child at an older age.** Women who give birth to their first child after age 30 may have an increased risk of breast cancer.
- **Having never been pregnant.** Women who have never been pregnant have a greater risk of breast cancer than do women who have had one or more pregnancies.
- **Postmenopausal hormone therapy.** Women who take hormone therapy medications that combine estrogen and progesterone to treat the signs and symptoms of menopause have an increased risk of breast cancer. The risk of breast cancer decreases when women stop taking these medications.
- **Drinking alcohol.** Drinking alcohol increases the risk of breast cancer.

Prevention

Breast cancer risk reduction for women with an average risk

Making changes in your daily life may help reduce your risk of breast cancer. Try to:

- **Ask your doctor about breast cancer screening.** Discuss with your doctor when to begin breast cancer screening exams and tests, such as clinical breast exams and mammograms.

  Talk to your doctor about the benefits and risks of screening. Together, you can decide what breast cancer screening strategies are right for you.

- **Become familiar with your breasts through breast self-exam for breast awareness.** Women may choose to become familiar with their breasts by occasionally inspecting their breasts during a breast self-exam for breast awareness. If there is a new change, lumps or other unusual signs in your breasts, talk to your doctor promptly.

  Breast awareness can't prevent breast cancer, but it may help you to better understand the normal changes that your breasts undergo and identify any unusual signs and symptoms.

- **Drink alcohol in moderation, if at all.** Limit the amount of alcohol you drink to no more than one drink a day, if you choose to drink.
- **Exercise most days of the week.** Aim for at least 30 minutes of exercise on most days of the week. If you haven't been active lately, ask your doctor whether it's OK and start slowly.
- **Limit postmenopausal hormone therapy.** Combination hormone therapy may increase the risk of breast cancer. Talk with your doctor about the benefits and risks of hormone therapy.

  Some women experience bothersome signs and symptoms during menopause and, for these women, the increased risk of breast cancer may be acceptable in order to relieve menopause signs and symptoms.

  To reduce the risk of breast cancer, use the lowest dose of hormone therapy possible for the shortest amount of time.

- **Maintain a healthy weight.** If your weight is healthy, work to maintain that weight. If you need to lose weight, ask your doctor about healthy strategies to accomplish this. Reduce the number of calories you eat each day and slowly increase the amount of exercise.
- **Choose a healthy diet.** Women who eat a Mediterranean diet supplemented with extra-virgin olive oil and mixed nuts may have a reduced risk of breast cancer. The Mediterranean diet focuses mostly on plant-based foods, such as fruits and vegetables, whole grains, legumes, and nuts. People who follow the Mediterranean diet choose healthy fats, such as olive oil, over butter and fish instead of red meat.

Breast cancer risk reduction for women with a high risk

If your doctor has assessed your family history and determined that you have other factors, such as a precancerous breast condition, that increase your risk of breast cancer, you may discuss options to reduce your risk, such as:

- **Preventive medications (chemoprevention).** Estrogen-blocking medications, such as selective estrogen receptor modulators and aromatase inhibitors, reduce the risk of breast cancer in women with a high risk of the disease.

  These medications carry a risk of side effects, so doctors reserve these medications for women who have a very high risk of breast cancer. Discuss the benefits and risks with your doctor.

- **Preventive surgery.** Women with a very high risk of breast cancer may choose to have their healthy breasts surgically removed (prophylactic mastectomy). They may also choose to have their healthy ovaries removed (prophylactic oophorectomy) to reduce the risk of both breast cancer and ovarian cancer.

CHAPTER 5: SYSTEM DESIGN

5.1 Machine Learning Model

5.1.1 Data Collection: I collect data from

Diabetes Dataset:

https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

Heart Disease Dataset:

https://archive.ics.uci.edu/ml/datasets/heart+disease

Parkinson's Dataset:

https://archive.ics.uci.edu/ml/datasets/Parkinson's

Breast Cancer Dataset:

https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)

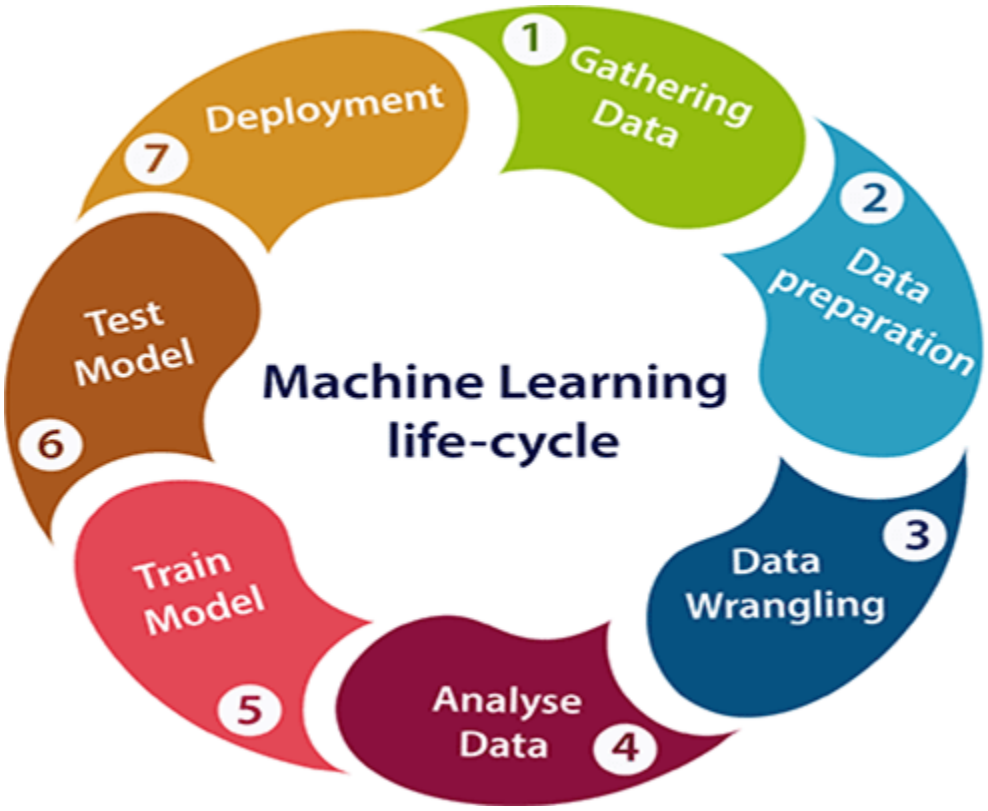

Fig -7: Machine Learning Life-Cycle

5.1.2 Diabetes Prediction Model :

```
df.head()
```

	Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

Table-2: Diabetes Data head()

Histogram

```
# Histogram and density graphs of all variables were accessed.
fig, ax = plt.subplots(4,2, figsize=(16,16))
sns.distplot(df.Age, bins = 20, ax=ax[0,0])
sns.distplot(df.Pregnancies, bins = 20, ax=ax[0,1])
sns.distplot(df.Glucose, bins = 20, ax=ax[1,0])
sns.distplot(df.BloodPressure, bins = 20, ax=ax[1,1])
sns.distplot(df.SkinThickness, bins = 20, ax=ax[2,0])
sns.distplot(df.Insulin, bins = 20, ax=ax[2,1])
sns.distplot(df.DiabetesPedigreeFunction, bins = 20, ax=ax[3,0])
sns.distplot(df.BMI, bins = 20, ax=ax[3,1])
```

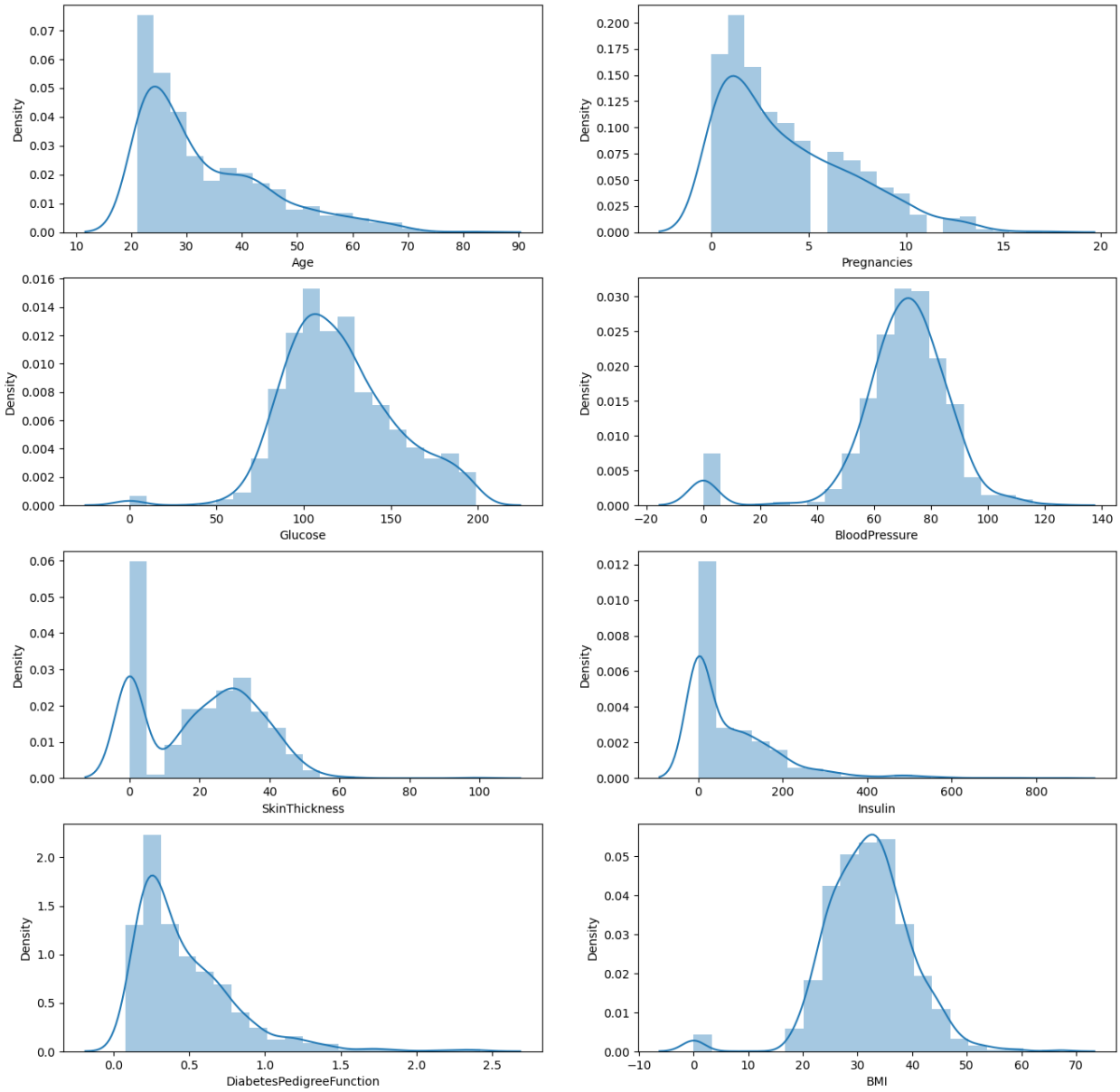

Fig-7: Diabetes Data Histogram

Algorithms

```
#Algorithms

def models(X_train, Y_train):
    # Logistic regression
    from sklearn.linear_model import LogisticRegression
    log=LogisticRegression(random_state=0)
    log.fit(X_train, Y_train)
```

```
# Decision Tree
from sklearn.tree import DecisionTreeClassifier
tree=DecisionTreeClassifier(random_state=0,criterion="entropy")
tree.fit(X_train, Y_train)

#random Forest
from sklearn.ensemble import RandomForestClassifier
forest=RandomForestClassifier(random_state=0, criterion="entropy",n_estimators=10)
forest.fit(X_train, Y_train)

#KNN
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 8)
classifier.fit(X_train, Y_train)

#Xgboost
from sklearn.ensemble import GradientBoostingClassifier
xgb = GradientBoostingClassifier(random_state = 12345)
xgb.fit(X_train, Y_train)

print('[0]logistic regression accuracy', log.score(X_train,Y_train))
print('[1]Decision Tree accuracy', tree.score(X_train,Y_train))
print('[2]random forest accuracy', forest.score(X_train,Y_train))
print('[3]KNN accuracy', classifier.score(X_train,Y_train))
print('[4]XGBoost accuracy', xgb.score(X_train,Y_train))

return log, tree, forest,classifier,xgb
```

model accuracy;

[0] logistic regression accuracy 0.7687296416938111

[1] Decision Tree accuracy 1.0

[2] random forest accuracy 0.995114006514658

[3] KNN accuracy 0.8566775244299675

[4] XGBoost accuracy 0.9788273615635179

5.1.3 Heart Disease Prediction Model:

```
df.head()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

Histogram:

```
sns.set_style('darkgrid')
df.hist(figsize=(30,30))
plt.show()
```

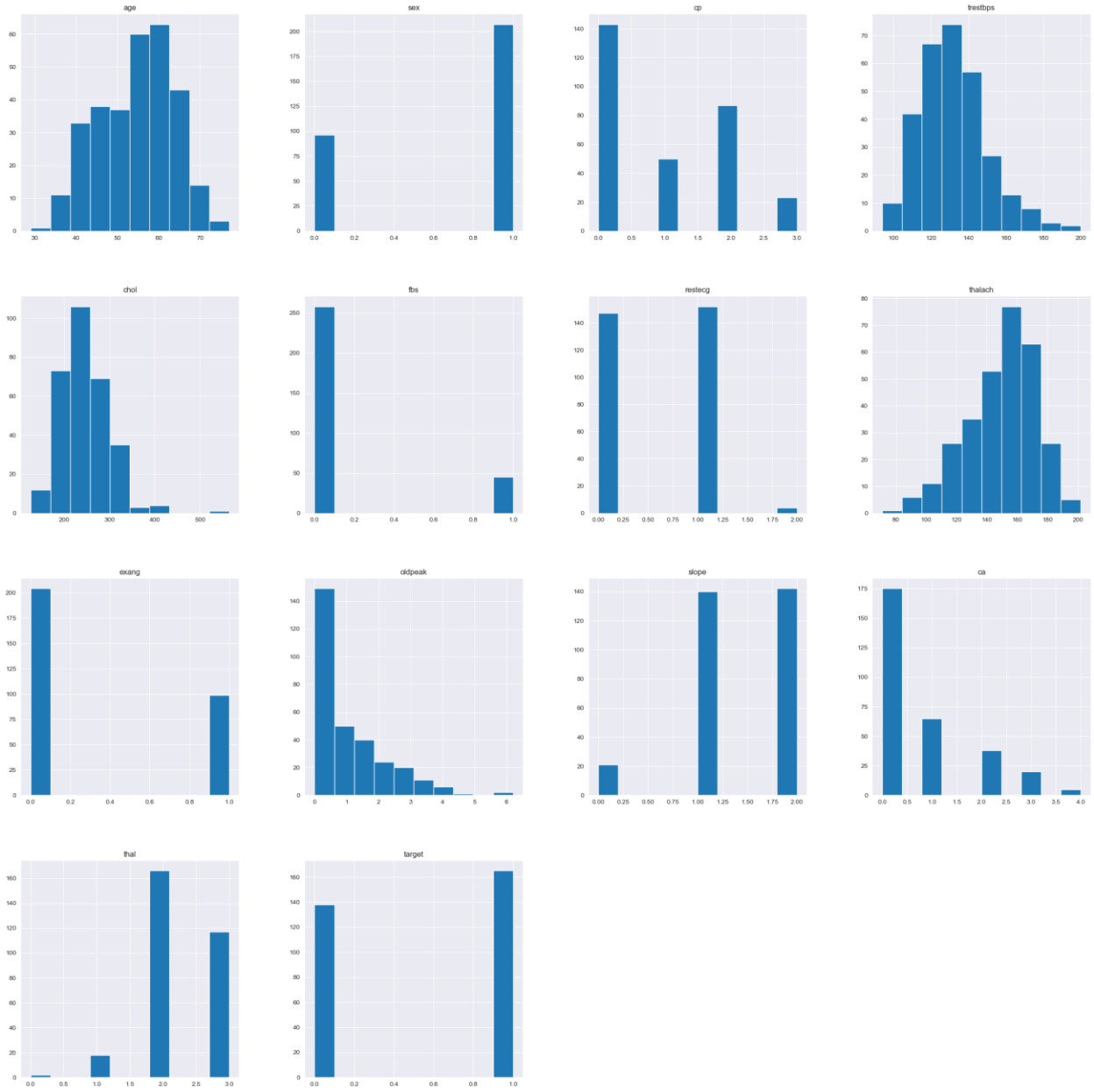

Fig-8: Heart disease data Histogram

Algorithms:

```
#Algorithms

def models(X_train, Y_train):
    # Logistic regression
    from sklearn.linear_model import LogisticRegression
    log=LogisticRegression(random_state=0)
    log.fit(X_train, Y_train)

    # Decision Tree
    from sklearn.tree import DecisionTreeClassifier
    tree=DecisionTreeClassifier(random_state=0,criterion="entropy")
    tree.fit(X_train, Y_train)

    #random Forest
    from sklearn.ensemble import RandomForestClassifier
    forest=RandomForestClassifier(random_state=0,
criterion="entropy",n_estimators=10)
    forest.fit(X_train, Y_train)

    #KNN
    from sklearn.neighbors import KNeighborsClassifier
    classifier = KNeighborsClassifier(n_neighbors = 8)
    classifier.fit(X_train, Y_train)
     #Xgboost
    from sklearn.ensemble import GradientBoostingClassifier
    xgb = GradientBoostingClassifier(random_state = 12345)
    xgb.fit(X_train, Y_train)
    #svm
    from sklearn import svm
    clf=svm.SVC(kernel='linear')
    clf.fit(X_train, Y_train)
    #CNN
    from sklearn.neural_network import MLPClassifier
    mlp = MLPClassifier(solver='lbfgs', alpha=1e-
5,hidden_layer_sizes=(2,50), random_state=1)
    mlp.fit(X_train, Y_train)

    print('[0]logistic regression accuracy', log.score(X_train,Y_train))
    print('[1]Decision Tree accuracy', tree.score(X_train,Y_train))
    print('[2]random forest accuracy', forest.score(X_train,Y_train))
    print('[3]KNN accuracy', classifier.score(X_train,Y_train))
    print('[4]XGBoost accuracy', xgb.score(X_train,Y_train))
```

```
    print('[5]SVM accuracy', clf.score(X_train,Y_train))
    print('[6]CNN accuracy', clf.score(X_train,Y_train))

    return log, tree, forest,classifier, xgb, clf, mlp
```

Accuracy:

[0]logistic regression accuracy 0.8429752066115702

[1]Decision Tree accuracy 1.0

[2]random forest accuracy 0.987603305785124

[3]KNN accuracy 0.8553719008264463

[4]XGBoost accuracy 1.0

[5]SVM accuracy 0.8429752066115702

[6]CNN accuracy 0.8429752066115702

5.1.4: Parkinson's Disease Prediction Model

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 24 columns):
 #   Column            Non-Null Count  Dtype  
---  ------            --------------  -----  
 0   name              195 non-null    object 
 1   MDVP:Fo(Hz)       195 non-null    float64
 2   MDVP:Fhi(Hz)      195 non-null    float64
 3   MDVP:Flo(Hz)      195 non-null    float64
 4   MDVP:Jitter(%)    195 non-null    float64
 5   MDVP:Jitter(Abs)  195 non-null    float64
 6   MDVP:RAP          195 non-null    float64
 7   MDVP:PPQ          195 non-null    float64
 8   Jitter:DDP        195 non-null    float64
 9   MDVP:Shimmer      195 non-null    float64
 10  MDVP:Shimmer(dB)  195 non-null    float64
 11  Shimmer:APQ3      195 non-null    float64
 12  Shimmer:APQ5      195 non-null    float64
 13  MDVP:APQ          195 non-null    float64
 14  Shimmer:DDA       195 non-null    float64
 15  NHR               195 non-null    float64
 16  HNR               195 non-null    float64
 17  status            195 non-null    int64  
 18  RPDE              195 non-null    float64
 19  DFA               195 non-null    float64
 20  spread1           195 non-null    float64
```

21 spread2 195 non-null float64

22 D2 195 non-null float64

23 PPE 195 non-null float64

dtypes: float64(22), int64(1), object(1)

memory usage: 36.7+ KB

Algorithms

```
#Algorithms

def models(X_train, Y_train):
    # Logistic regression
    from sklearn.linear_model import LogisticRegression
    log=LogisticRegression(random_state=0)
    log.fit(X_train, Y_train)

    # Decision Tree
    from sklearn.tree import DecisionTreeClassifier
    tree=DecisionTreeClassifier(random_state=0,criterion="entropy")
    tree.fit(X_train, Y_train)

    #random Forest
    from sklearn.ensemble import RandomForestClassifier
    forest=RandomForestClassifier(random_state=0,
criterion="entropy",n_estimators=10)
    forest.fit(X_train, Y_train)

    #KNN
    from sklearn.neighbors import KNeighborsClassifier
    classifier = KNeighborsClassifier(n_neighbors = 8)
    classifier.fit(X_train, Y_train)
     #Xgboost
    from sklearn.ensemble import GradientBoostingClassifier
    xgb = GradientBoostingClassifier(random_state = 12345)
    xgb.fit(X_train, Y_train)
    #svm
    from sklearn import svm
    clf=svm.SVC(kernel='linear')
```

```
    clf.fit(X_train, Y_train)
    #CNN
    from sklearn.neural_network import MLPClassifier
    mlp = MLPClassifier(solver='lbfgs', alpha=1e-
5,hidden_layer_sizes=(2,50), random_state=1)
    mlp.fit(X_train, Y_train)


    print('[0]logistic regression accuracy', log.score(X_train,Y_train))
    print('[1]Decision Tree accuracy', tree.score(X_train,Y_train))
    print('[2]random forest accuracy', forest.score(X_train,Y_train))
    print('[3]KNN accuracy', classifier.score(X_train,Y_train))
    print('[4]XGBoost accuracy', xgb.score(X_train,Y_train))
    print('[5]SVM accuracy', clf.score(X_train,Y_train))
    print('[6]CNN accuracy', mlp.score(X_train,Y_train))

    return log, tree, forest,classifier, xgb, clf, mlp
```

Accuracy:

[0] logistic regression accuracy 0.8782051282051282

[1]Decision Tree accuracy 1.0

[2] random forest accuracy 1.0

[3]KNN accuracy 0.9487179487179487

[4] XGBoost accuracy 1.0

[5]SVM accuracy 0.8974358974358975

[6]CNN accuracy 0.9487179487179487

5.1.5 Breast Cancer Prediction Model

```
# Plot histograms for each variable
sns.set_style('darkgrid')
df.hist(figsize=(30,30))
plt.show()
```

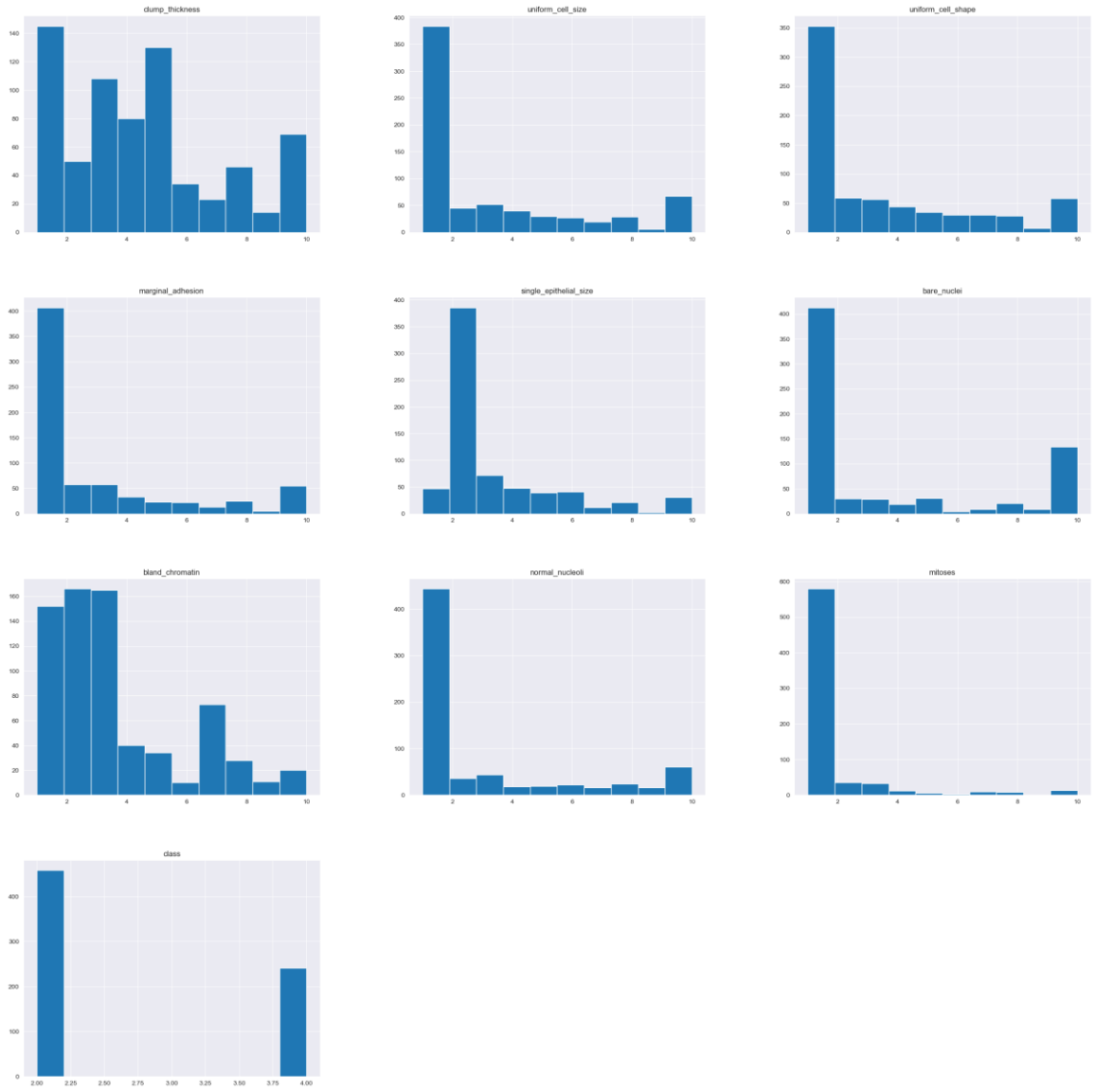

Fig-9: Breast Cancer Data Histogram

Algorithms:

```
# Define models to train
models= []
models.append(('CART', DecisionTreeClassifier()))
models.append(('SVM', SVC()))
models.append(('NB', GaussianNB()))
models.append(('KNN', KNeighborsClassifier()))

# evaluate each model in turn
results = []
names = []

for name, model in models:
    kfold = KFold(n_splits=10)
    cv_results = cross_val_score(model, X_train, Y_train, cv=kfold,
scoring=scoring)
    results.append(cv_results)
    names.append(name)
    msg = "For %s Model:Mean accuracy is %f (Std accuracy is %f)" % (name,
cv_results.mean(), cv_results.std())
    print(msg)
```

Accuracy;

For CART Model: Mean accuracy is 0.956973 (Std accuracy is 0.025180)

For SVM Model: Mean accuracy is 0.971386 (Std accuracy is 0.013512)

For NB Model: Mean accuracy is 0.963223 (Std accuracy is 0.025463)

For KNN Model: Mean accuracy is 0.969345 (Std accuracy is 0.016428)

7.3 User Interface

User Interface:

Stream-lit App:

Home Page

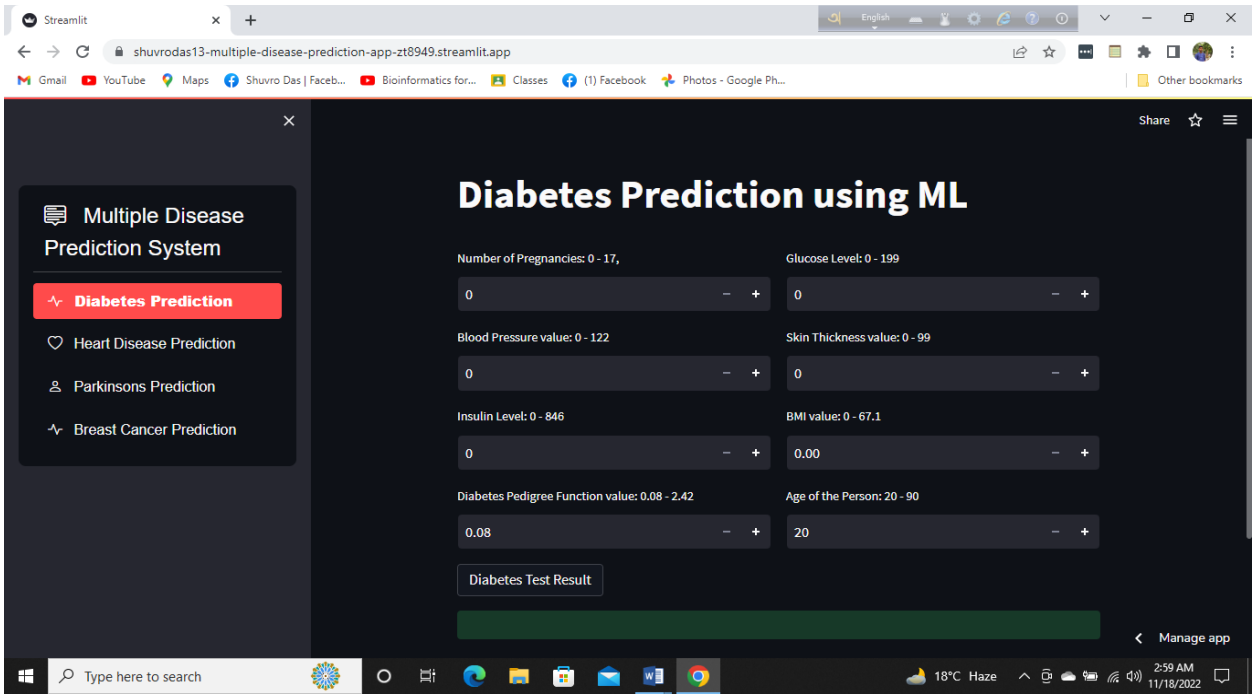

Fig-10: Home page

Diabetes Prediction:

Positive:

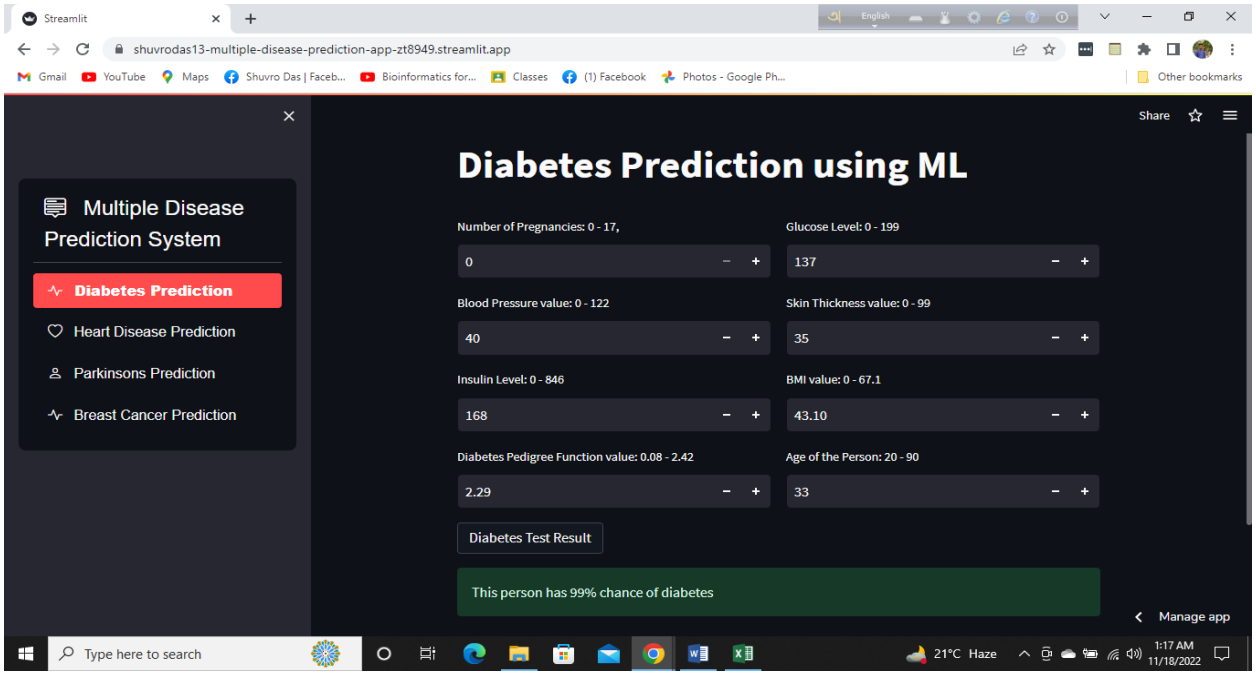

Fig:11: Diabetes Positive

Negative:

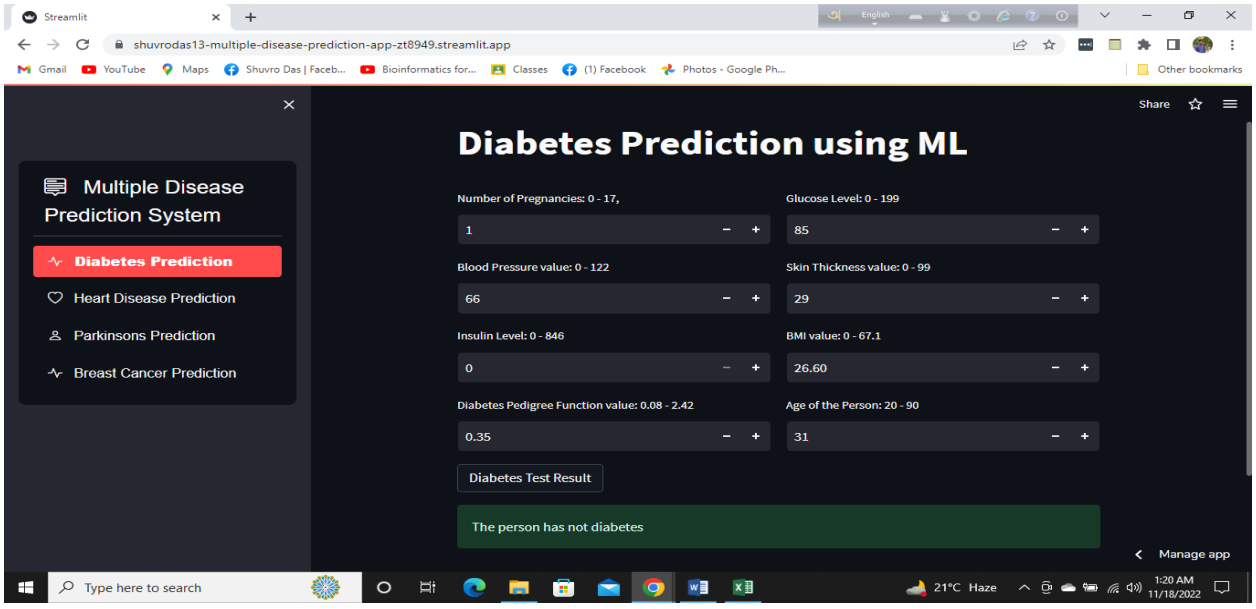

Fig-12: Diabetes negative

Heart Disease Prediction

Positive:

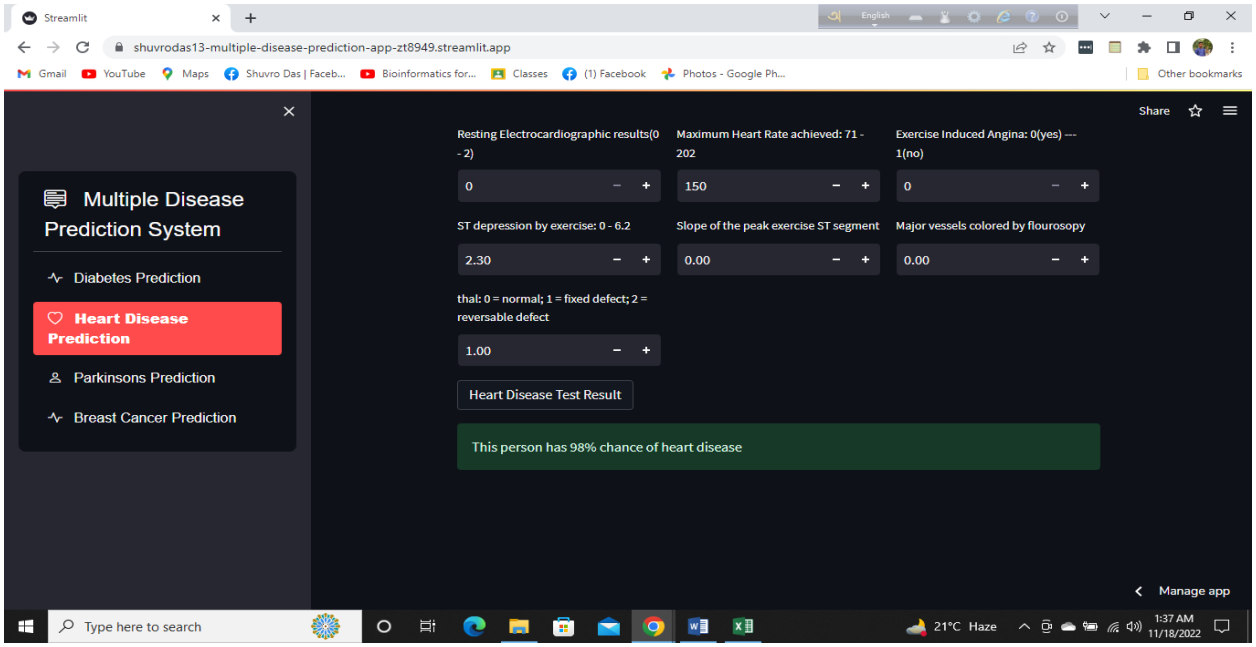

Fig-13: Heart Disease Positive

Negative:

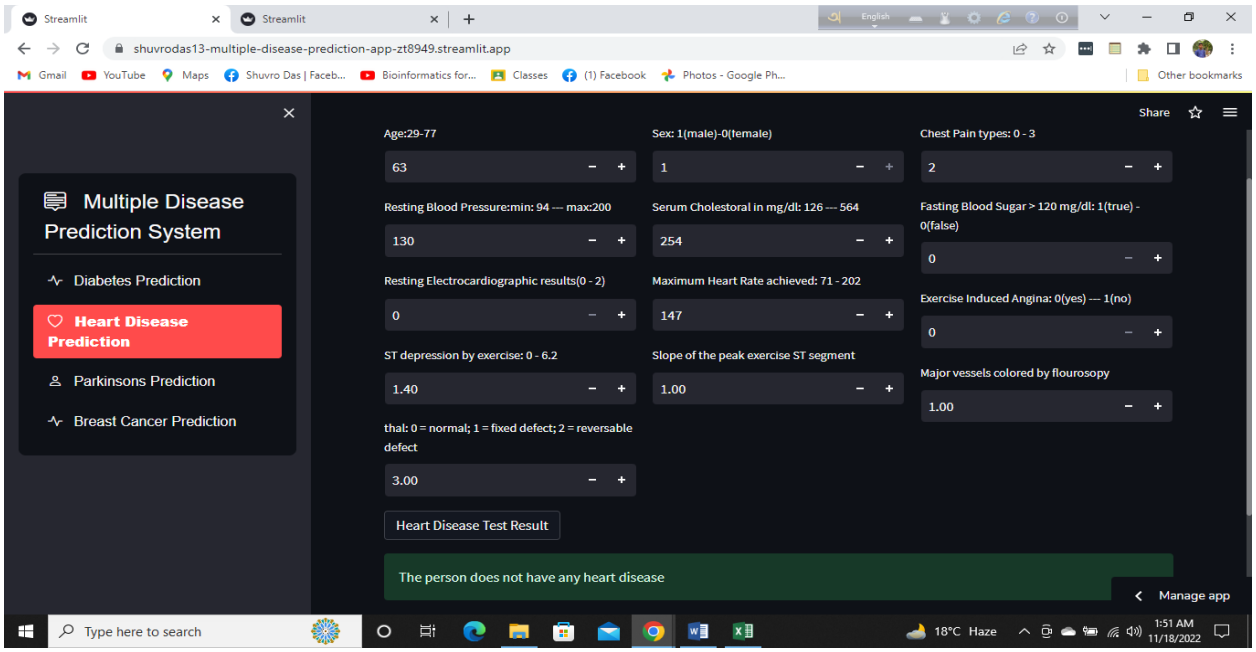

Fig-14: Heart Disease Negative

Parkinson's Disease Prediction

Positive

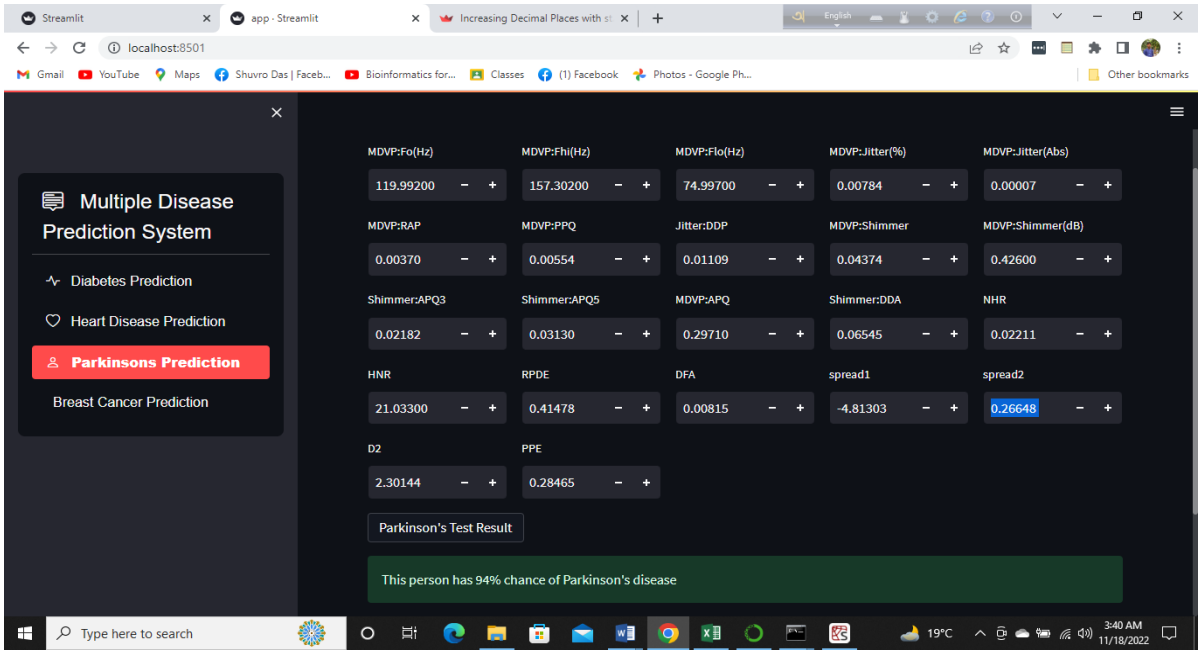

Fig-15: Parkinson's disease Positive

Negative

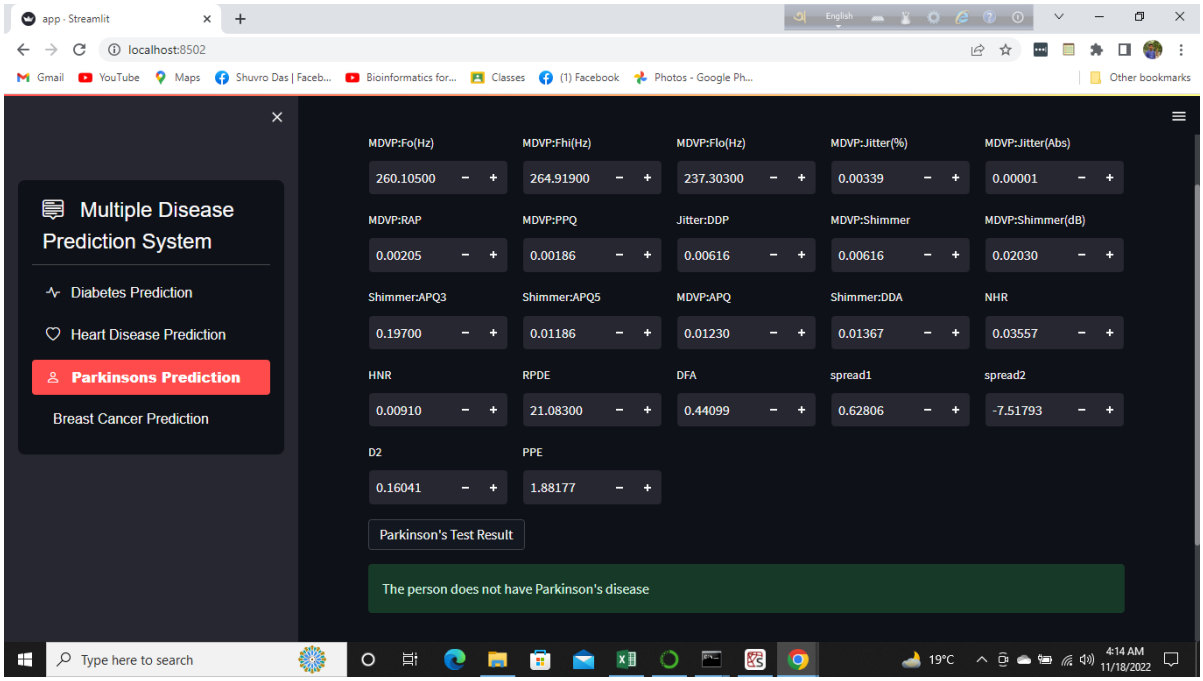

Fig-16: Parkinson's disease Negative

Breast Cancer Prediction

Positive

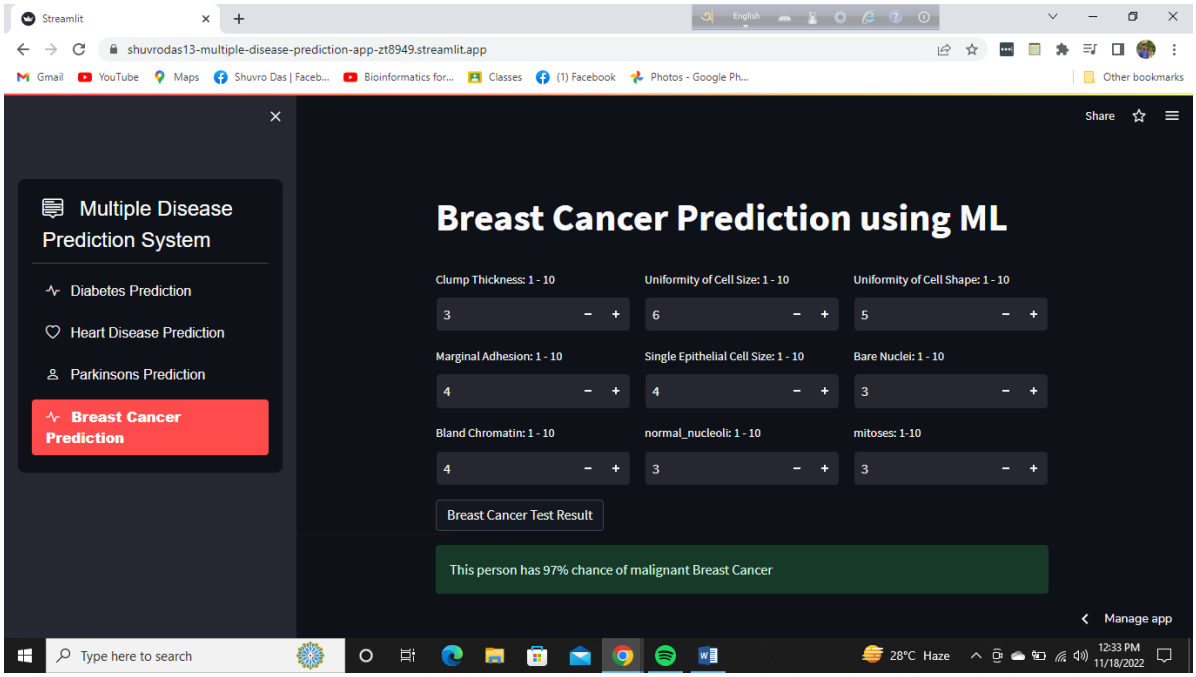

Fig-17: Breast Cancer Positive

Negative

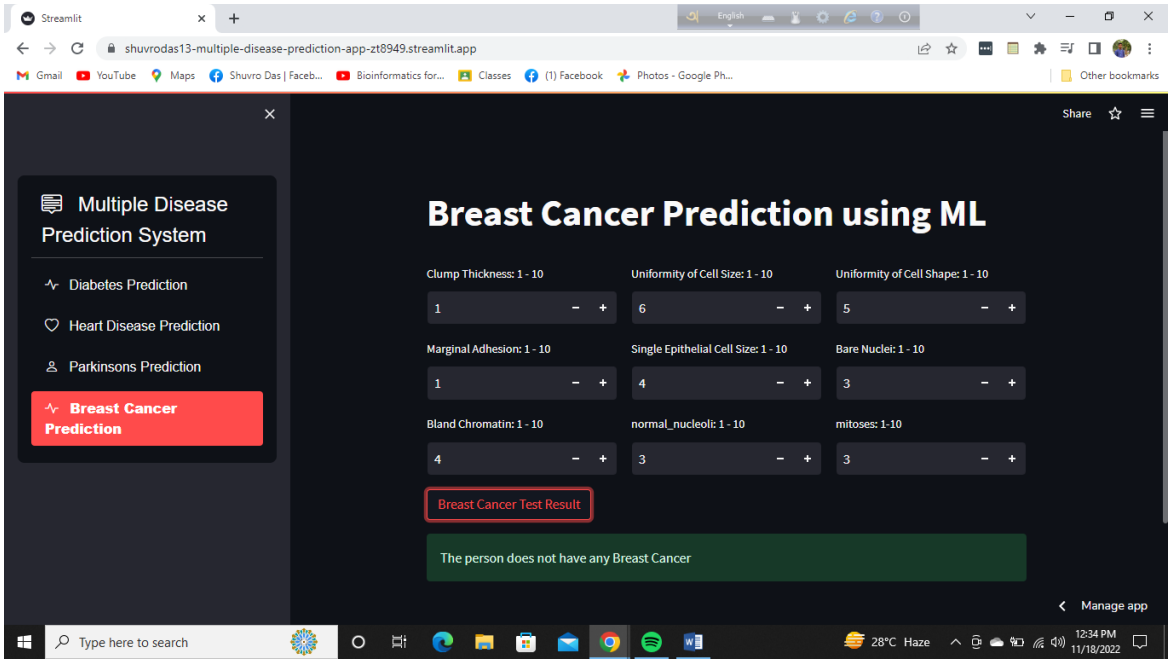

Fig-18: Breast Cancer Negative

CHAPTER 6: CONCLUSION

In this project I have used machine learning algorithms to detect accuracy of different disease prediction model and Stream-lit app for user interface. Diabetes, Heart Disease, Parkinson's disease and Breast Cancer can be predicted by Multiple Disease Prediction System.

In Diabetes Prediction There is 99% accuracy in prediction. In Heart disease Prediction there is 98% accuracy in prediction. In Parkinson's disease Prediction there is 94% accuracy in prediction. In Breast Cancer Prediction there is 97% accuracy in prediction. Because of this project the user doesn't need to traverse different websites which saves time as well. Diseases if predicted early can increase your life expectancy as well as save you from financial troubles. Multiple disease prediction model is used to predict multiple diseases at a time. Here based on the user input disease will be predicted. The choice will be given to the user. If the user wants to predict a particular disease or if the user doesn't enter any disease type then based on user entered inputs corresponding disease model will be invoked and predicted. The advantage of a multiple disease prediction model in advance can predict the probability of occurrence of various diseases and also can reduce mortality ratio.

CHAPTER 7: FUTURE WORK

I want to enhance this application in future. I wish to add multiple features and functions in this app. So that user can easily identify different diseases and their cure procedure. In the future we can add more diseases in the existing API. We can try to improve the accuracy of prediction in order to decrease the mortality rate. Try to make the system user-friendly and provide a chat bot for normal queries. In that Project, I've used Random Forest, XG Boost, Naive Bayes, KNN, CNN, Logistic Regressions and others algorithms. In future I like to work along with Image Processing, Language Processing, Neural Network and Deep Learning. I also wish to work on other diseases that may help to survive the mankind world-wide.

References

1. S. Sharma, A. Aggarwal and T. Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 2018, pp. 114-118, doi: 10.1109/CTEMS.2018.8769187.
2. Shah, D., Patel, S. & Bharti, S.K. Heart Disease Prediction using Machine Learning Techniques. *SN COMPUT. SCI.* 1, 345 (2020). https://doi.org/10.1007/s42979-020-00365-y.
3. Belić M, Bobić V, Badža M, Šolaja N, Đurić-Jovičić M, Kostić VS. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease-A review. Clin Neurol Neurosurg. 2019 Sep;184:105442. doi: 10.1016/j.clineuro.2019.105442. Epub 2019.
4. N. Abdulhadi and A. Al-Mousa, "Diabetes Detection Using Machine Learning Classification Methods," 2021 International Conference on Information Technology (ICIT), 2021, pp. 350-354, doi: 10.1109/ICIT52682.2021.9491788.
5. **Breast Cancer Details**
   Source:https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470?p=1
6. **Parkinson's Disease Details**
   Source: https://www.mayoclinic.org/diseases-conditions/Parkinson's-disease/symptoms-causes/syc-20376055
7. **Diabetes Details**
   Source: https://www.who.int/news-room/fact-sheets/detail/diabetes
8. **Heart Disease Details**

   Source: https://www.mayoclinic.org/diseases-conditions/heart-

   disease/symptoms-causes/syc-20353118