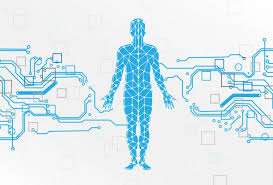
Disease Prediction System for Parkinson Disease and Breast Cancer using Voice Command

**[Project Documentation]**

**SUBMITTED FOR PARTIAL FULFILMENT OF**

**MASTER OF COMPUTER APPLICATIONS**

**[2021 – 2023]**

**Ashutosh Kumar Yadav 504121011006**

**Mujahid Ali Ansari 504121011025**

**Shahil Kumar Chourasia 504121011045**

**Sirsha Majumder 504121021050**

**Sonu Routh 504121011051**

**UNDER THE GUIDANCE OF DR. ANANJAN MAITY**

**(ASSISTANT PROFESSOR OF GURUNANAK INSTITUTE OF TECHNOLOGY)**

**Logo, company name  Description automatically generatedDEPARTMENT OF COMPUTER APPLICATIONS GURUNANAK INSTITUTE OF TECHNOLOGY**

|  |  |  |
| --- | --- | --- |
| **SL.NO** | **CONTENT TOPICS** | **PAGE NUMBER** |
| **1** | **CERTIFICATE OF APPROVAL** | 3 |
| **2** | **ACKNOWLEDGEMENT** | 4 |
| **3** | **ABSTRACT** | 5 |
| **4** | **INTRODUCTION** | 6 |
| **5** | **OBJECTIVE** | 7 |
| **6** | **FLOW CHART DIAGRAM** | 8 |
| **7** | **TOOLS AND TECHNOLOGY** | 9 |
| **8** | **APPROACH** | 10-24 |
| **9** | **OUTPUTS** | 25-29 |
| **10** | **EDA OF PARKINSON AND BREAST CANCER DISEASE** | 30-37 |
| **11** | **PROPOSED METHODOLOGIES** | 38-43 |
| **12** | **ACCURACY TABLE** | 44 |
| **13** | **APPLICATIONS** | 45 |
| **14** | **STRENGTHS** | 46 |
| **15** | **LIMITATIONS** | 47 |
| **16** | **CONCLUSION** | 48 |
| **17** | **FUTURE DEVELOPMENT** | 49 |
| **18** | **REFERENCES** | 50 |
| **19** | **INDIVIDUAL CONTRIBUTION** | 51 |

## **TABLE OF CONTENTS**

**CERTIFICATE OF APPROVAL**

This is to certify that Project work entitled “Self Health Monitoring System for Parkinson Disease and Breast Cancer Prediction Using Voice Command” Carried out by Ashutosh Kumar Yadav 504121011006, Mujahid Ali Ansari 504121011025, Shahil Kumar Chourasia 504121011045, Sirsha Majumder 504121021050, Sonu Routh 504121011051.under the   
Mentorship of Prof. Ms. Dola Saha and the wonderful guidance of Prof. Dr. Anjan Maity and Prof Mr. Chiranjib Dutta(H.O.D) of 4th Semester Master of Computer Applications from Gurunanak Institute of Technology, Kolkata 700114 from West Bengal has been satisfactorily completed by them and is thus worthy of acceptance for the degree of   
Master of Computer Applications.



**ACKNOWLEDGEMENT**

We, all the tam members, would like to take this opportunity to express our profound gratitude and kind regards to our guide **Dr. Ananjan Maity** (Assistant Professor – CA) for his exemplary guidance, monitoring, and constant encouragement throughout this project work. The blessings, help and guidance given by him to us shall carry us to a long way in our journeys of life on which we are about to embark.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Signature of Students Signature of Mentor**

**ABSTRACT**

This project introduces a System that can be used to predict if a patient is affected by **Parkinson Disease** or by **Breast Cancer**. Even today, these diseases are very serious health issues that are persisting.

This system is quite ready to be used by the administration of hospitals as well – the staffs can use it to register new patients and predict if they have symptoms of Parkinson disease or Breast Cancer, send the patients’ health reports to their respective mail ids, retrieve health reports of existing patients in database and to discharge patients. Having all these registrations, predictions, etc. done before getting checked by the main doctor cuts off time consumption and makes check-ups easy for doctors as well.

Also, this system can be used by the Staff Head(admin) to access the data of all the working staffs from the database along with all the patients’ details and delete staff data if required. The admin only would have a very confidential password for this special access.

The add-on feature of this system is that, we **can use voice commands** to get the actions of the system done. This system is available to be used 24X7 by staffs and admin.

**INTRODUCTION**

Immediate treatment and accurate diagnosis of neurological disorders is critical for care and research. However, in case of Parkinson Disease (PD) needs special care and treatment to reduce the disability of patients and cope up with changes in their physical capabilities because of this progressive neurodegenerative disorder and for people affected by Breast Cancer (BC) to have an early detection of this disease that could help slow down the progress of the disease and reduce the mortality rate through appropriate therapeutic interventions at the right time.

Parkinson disease (PD) is the most common type of parkinsonism and the second most common progressive neurodegenerative disorder. It affects about 1% of the population over age 50 years and about 2.5% of the population over age 70. The lifetime risk for PD development is 2.0% in men and 1.3% in women. Idiopathic, also known as sporadic PD, is the most common form of PD, affecting primarily older adults. In general, PD is associated with motor symptoms, such as resting tremor, bradykinesia/akinesia, and rigidity because of dopamine deficiency in the basal ganglia due to neurodegeneration of dopaminergic neurons in the substantia nigra *pars compacta* (SNPC). In addition to motor symptoms, nonmotor symptoms and complications, such as neuropsychiatric or neurobehavioral problems, autonomic dysfunction, and sensory problems, are also considered an important part of PD. The disease Cancer starts when normal cell in the breast grow abnormally and out of control, leading to a tumour, which can be malignant or benign. A malignant tumour can grow and spread affecting other body part whereas a benign tumour can grow but does not spread.

A tumour can be cancerous or benign. A cancerous tumour is malignant, meaning it can grow and spread to other parts of the body. A benign tumour means the tumour can grow but has not spread If breast cancer comes back after initial treatment, it can recur locally, meaning in the same breast and/or regional lymph nodes. It can also recur elsewhere in the body, called a [distant recurrence or metastatic recurrence](https://www.cancer.net/node/36246).

Our system has the potential to predict the early onsets of these vital diseases which might give enough time to get it cured or prevent malignancy.

**OBJECTIVE**

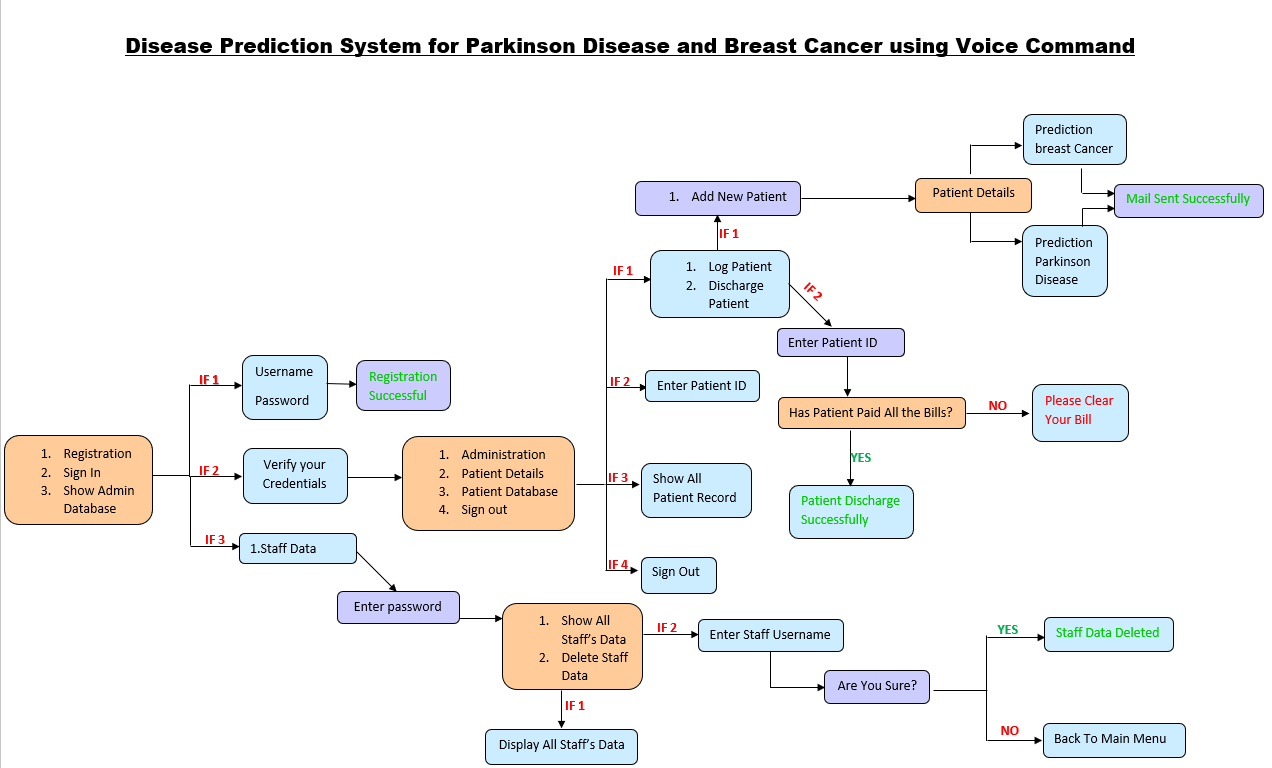
The primary aim of our project this project is to **predict the Parkinson and Breast Cancer diseases** from the given symptoms by patients and create and monitor a health profile of every individual patient.

**Health is the most important factor for everyone**. But unfortunately it has been neglected today for many reasons. Absence of doctor due to some reason during emergency may result in loss of life. Not only that sometimes patient often feel hesitant to go to hospital for minor symptoms. These may prompt into major illness. With the proliferation of technology in health care becomes easier to diagnosis any disease – even the deadliest ones. It is applied in healthcare to identify the clusters of patients, diseases, and future predictions using different machine learning tools. So, this project work proposed a Disease Prediction System that will help user to receive immediate guidance regarding their health issues.

To diagnose any disease, doctor initially analyses the symptoms of the patient and after that the result is predicted. Similarly, machine diagnose the diseases based on the symptoms just like the doctor does. The system is fetched with various symptoms and their disease related with it.

This system **aims to improve disease treatment and its diagnosis in early stages for a faster and better treatment**. Therefore, it is an attempt to make a faster and more accurate disease prediction and **help the physicians for making a reliable decision in a short span of time by increasing efficiency and quality in health management system**.

**FLOW CHART DIAGRAM**

****

The above flow-chart diagram represents the total workflow of the hospital management system including all operations of patients’ database as well as staffs’ database and predicts if the patient is suffering from Parkinson Disease or Breast Cancer or any other disease.

**TOOLS AND TECHNOLOGY**

**Tools and Software**

* Visual Studio Code:

Visual Studio Code is a code editor that is redefined and optimized for building and debugging modern web and cloud applications. Visual it is free and also available on our favourite platforms – Linux, macOS, and Windows.

* MySql :

MySQL Workbench is a graphical tool that works with [MySQL](https://database.guide/what-is-mysql/). MySQL Workbench offers an easy-to-use interface to perform multiple tasks involved when working with databases. It integrates SQL development, administration, database design, creation, and maintenance into one visual integrated development environment. MySQL Workbench is similar to SQL Server’s [SSMS](https://database.guide/what-is-sql-server-management-studio/), which is used for administering [SQL Server](https://database.guide/what-is-sql-server/).

**Programming Language :**

* Python :

Python is a high-level programming language, that precludes the need to compile code before executing a program because Python does the compilation in the background. Because Python is a high-level programming language, it abstracts many sophisticated details from the programming code.

**Skill-based Technology :**

* Machine Learning :

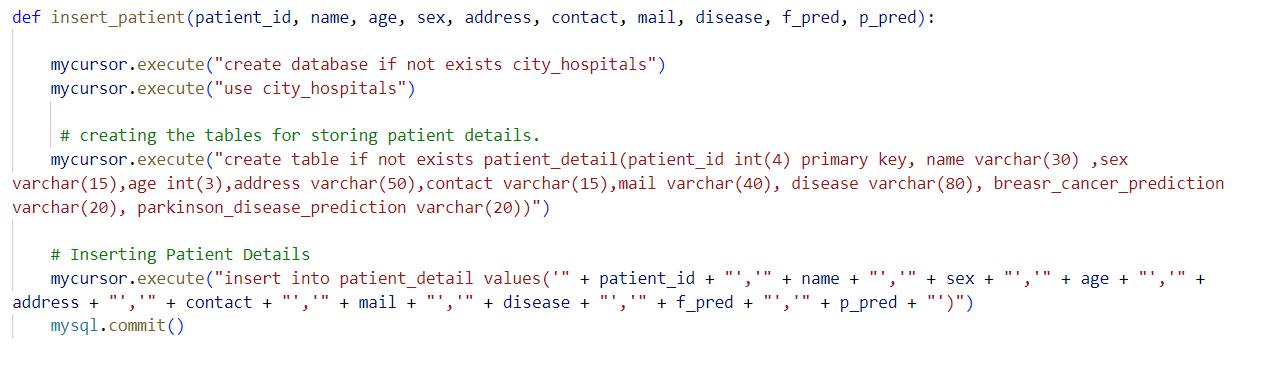
Machine learning is a subfield of artificial intelligence which includes the development of algorithms and statistical models that enable computers to improve their performance in tasks through experience. These algorithms and models are designed to learn from data and make predictions or decisions without explicit instructions.

**APPROACH**

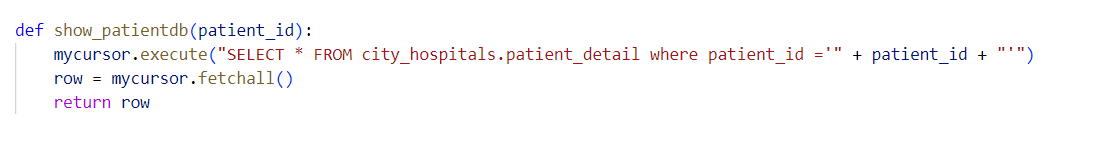
* **Database**
* Patient\_db.py
* insert\_patient()
* show\_patientdb()
* delete\_patientdb()
* show\_all\_patientdb()
* Staff\_db.py
* insert\_staff()
* show\_staff()
* show\_all\_staffdb()
* delete\_staff()
* **Predicted Disease**
* Parkinson.py
* Parkinson()
* Breast\_cancer.py
* Breast\_cancer()
* **Send Mail**
* Email.py
* PDF()
* Send\_mail()
* **Voice Command**
* Command.py
* takeCommand()
* Speak.py
* Speak()
* Wish.py
* wishMe()
* **main.py**

1. **Database**

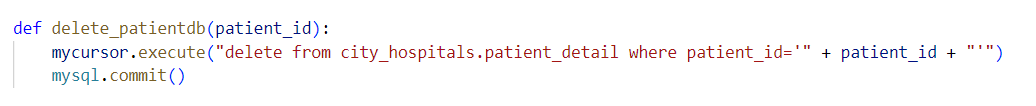
* **Patient\_db.py**
* **insert\_patient()** : This function is used to insert patient details in a database, especially when a new patient is being registered.



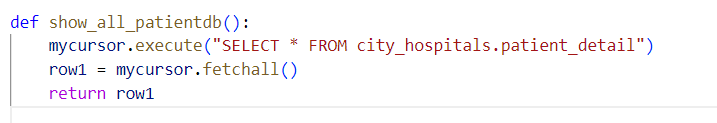
* **show\_patientdb()** : This function is used to retrieve patients’ records. Staff members can get specific patient details by using their patient ID to keep the track of their health issues.



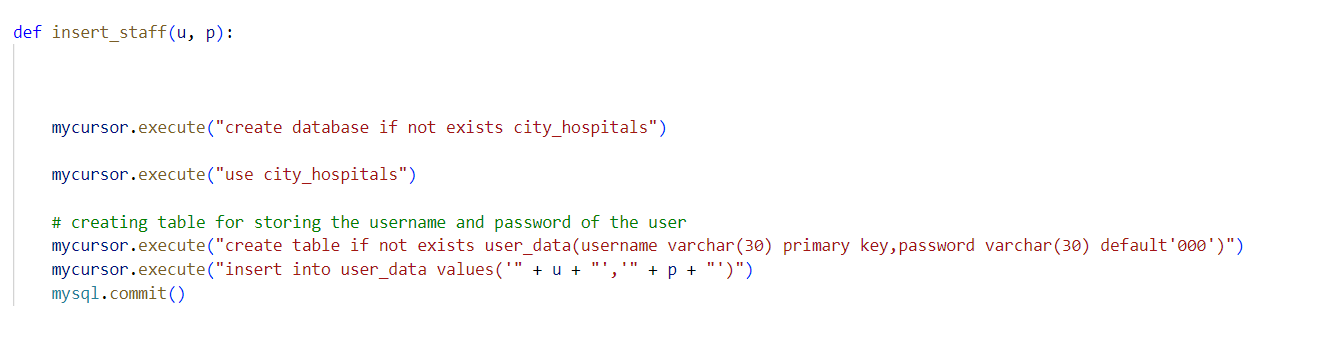
* **delete\_patientdb()** : This function is used to delete the patient details from the database if it’s no longer required.



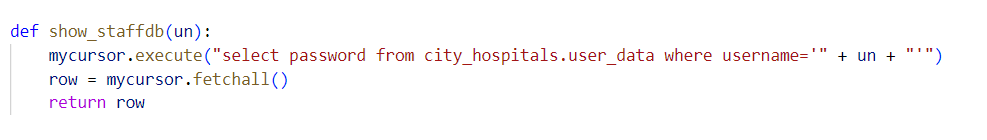
* **show\_all\_patientdb()** : This function is used to show all the patients’ records present in the database.



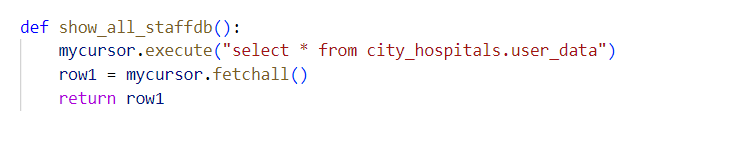
* **Staff\_db.py**
* **insert\_staff()** : This function is used to insert staff details in the database like their username and password.



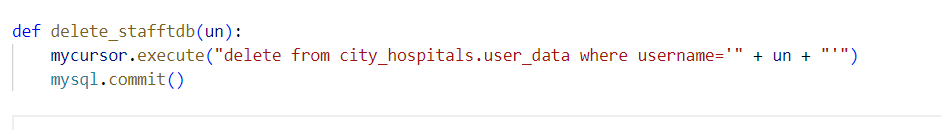
* **show\_staff()** : This function is used to retrieve all the Staff details from the database.

****

* **show\_all\_staffdb()** : This function is used to show all the staff records present in the database.

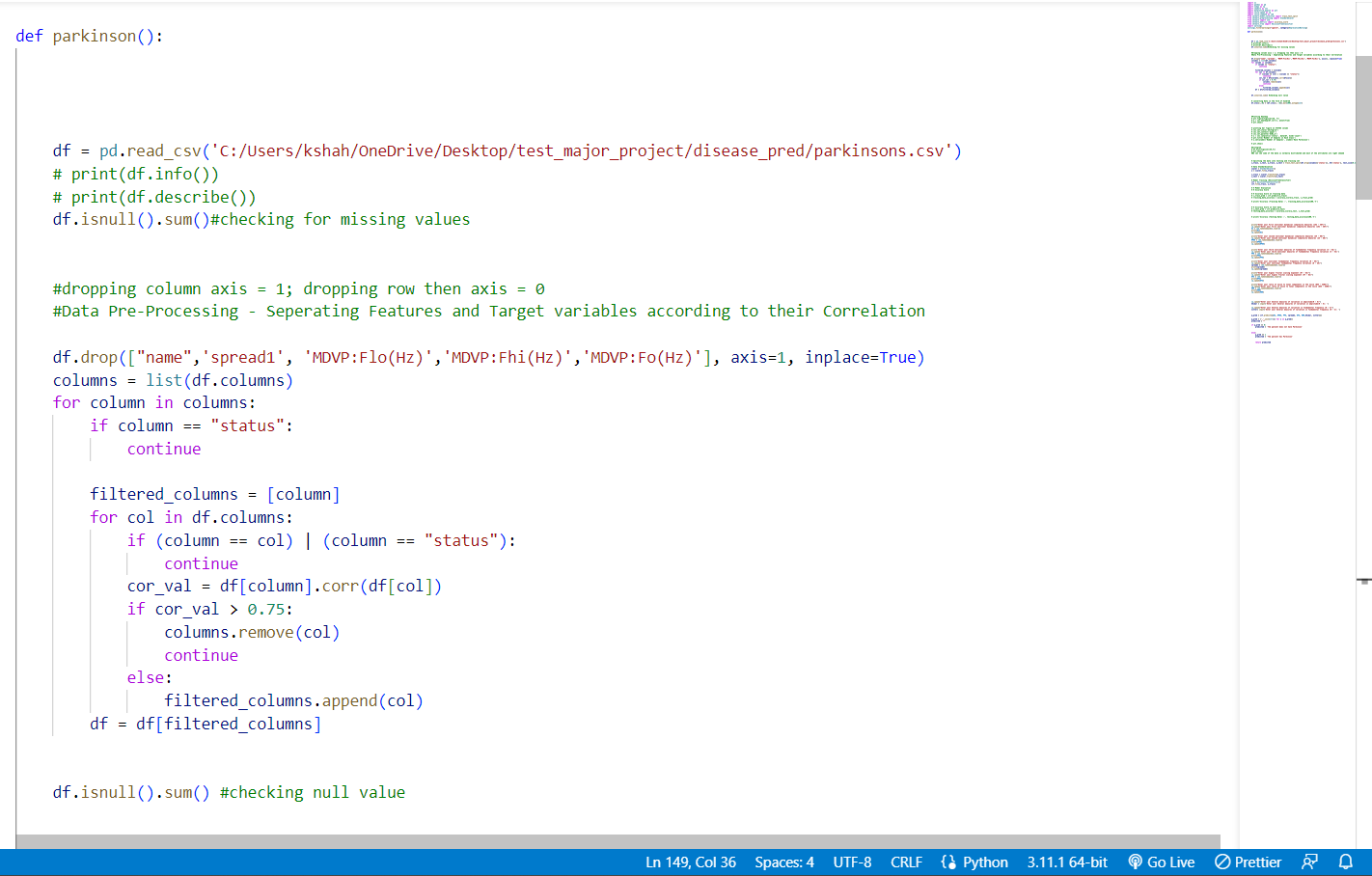


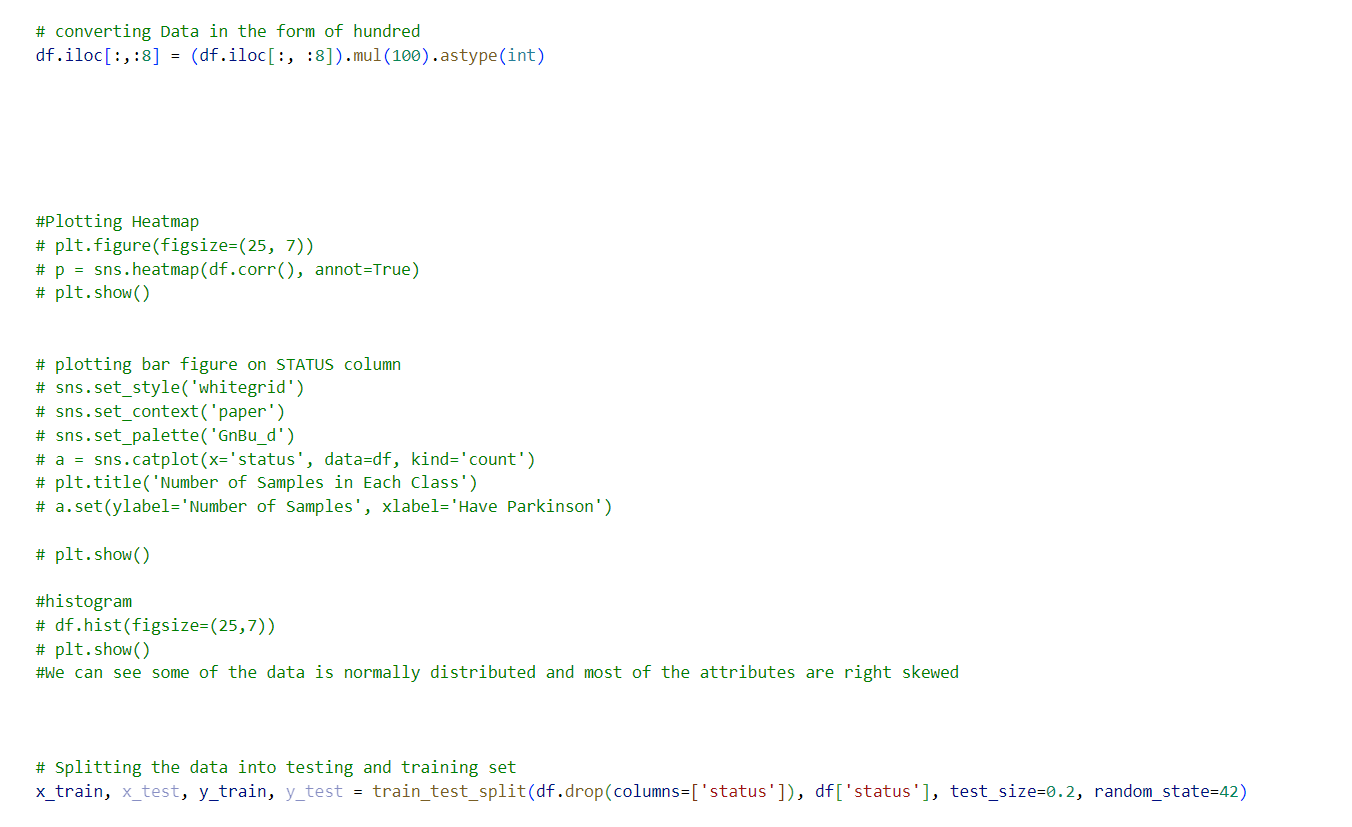
* **delete\_staff()** : This function is used to delete the staff details from the database if it’s no longer required.

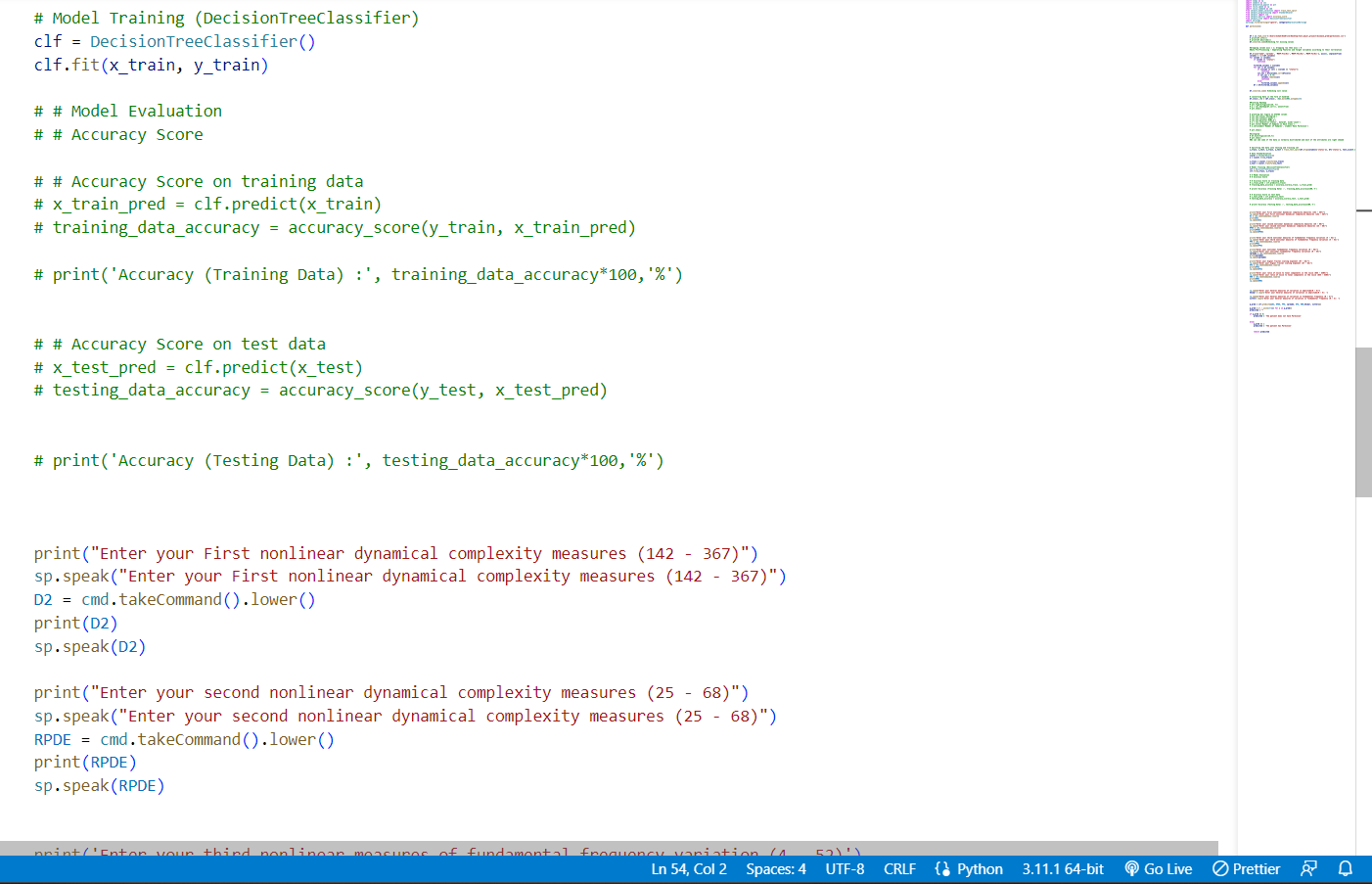
****

1. **Predicted Disease**

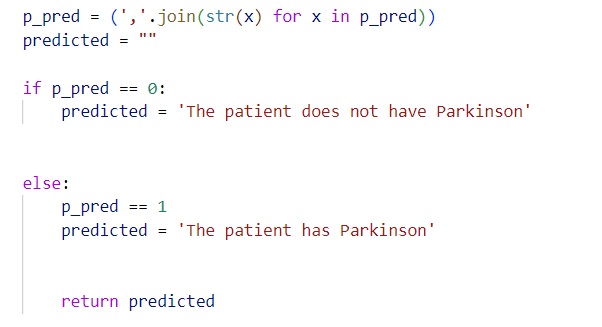
* **Parkinson.py**
* **Parkinson()** : This function is used to predict the Parkinson Disease.



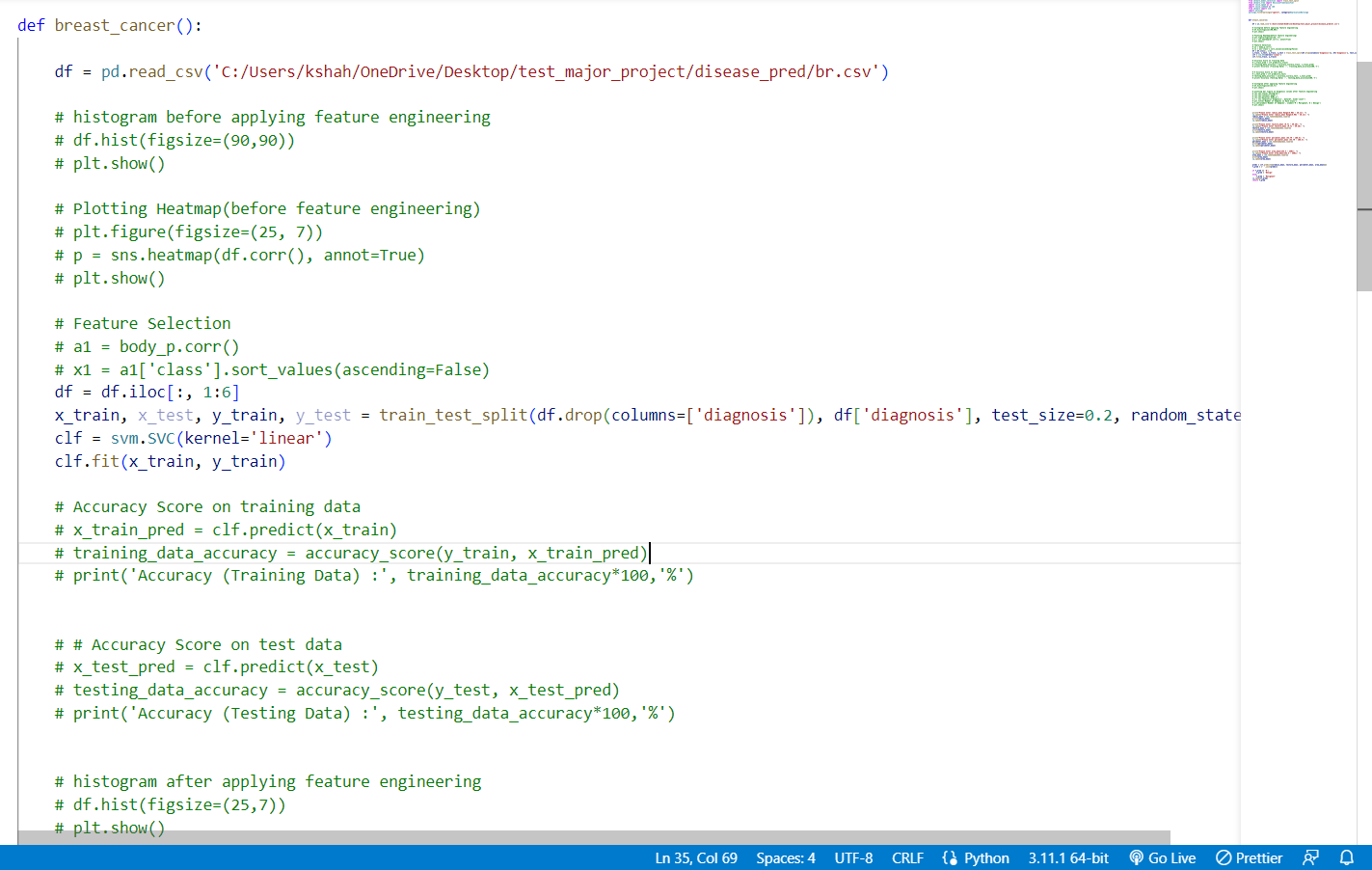


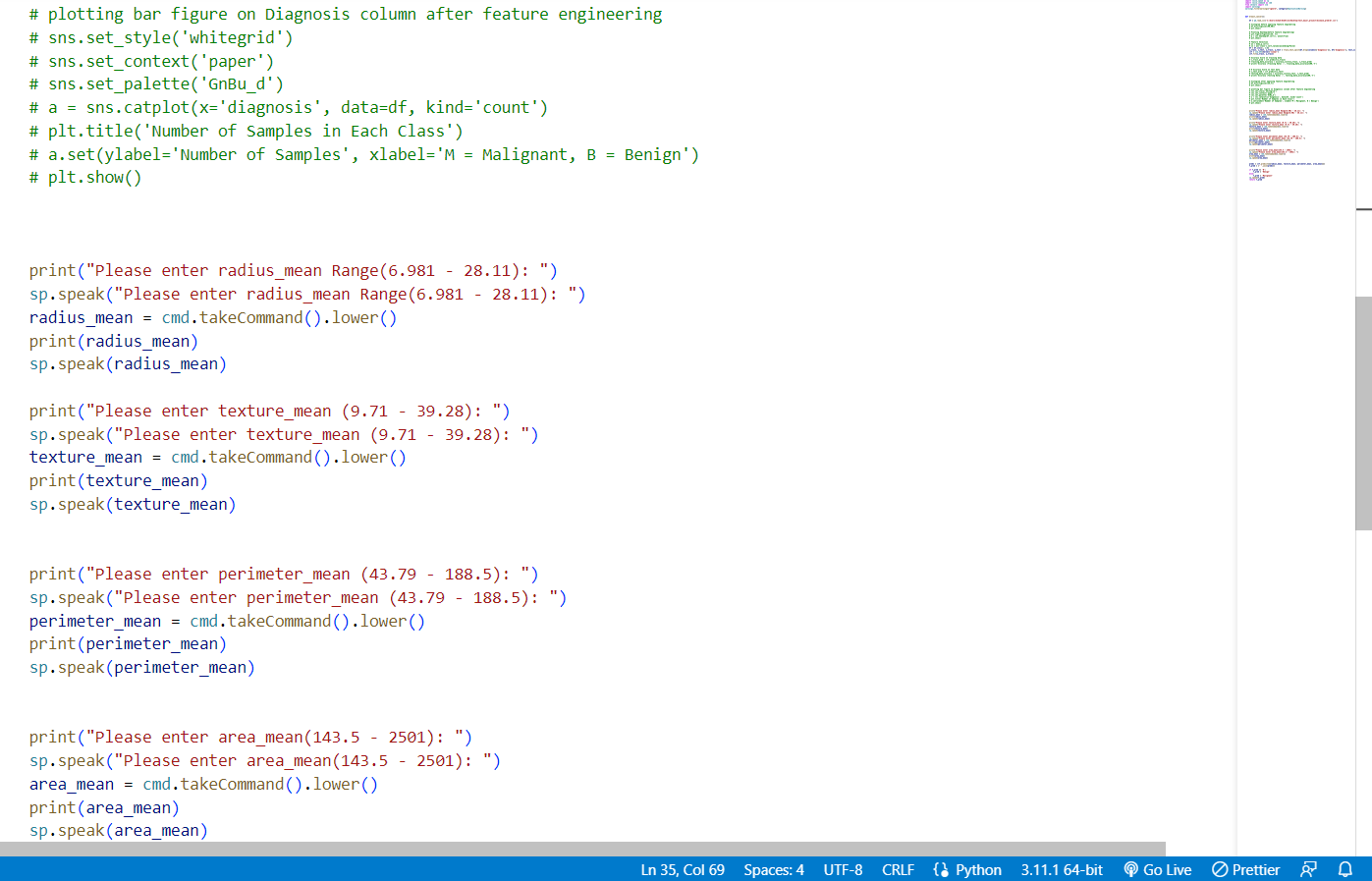


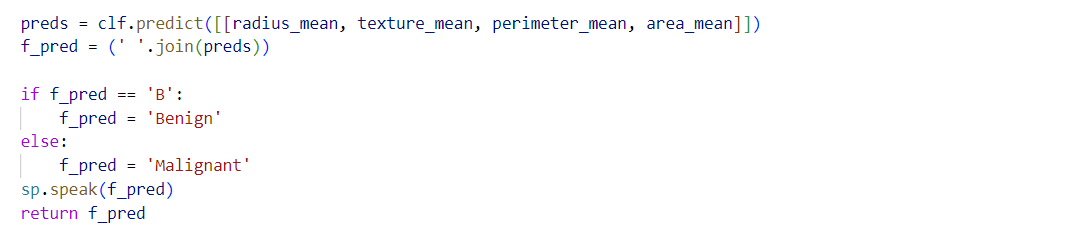




* **Breast\_cancer.py**
* **Breast\_cancer()** : This function predicts the type of Breast Cancer.

****

****

****

1. **Send Mail**

* **Email.py**
* **PDF()** : This function is used to create patient reports in pdf format.



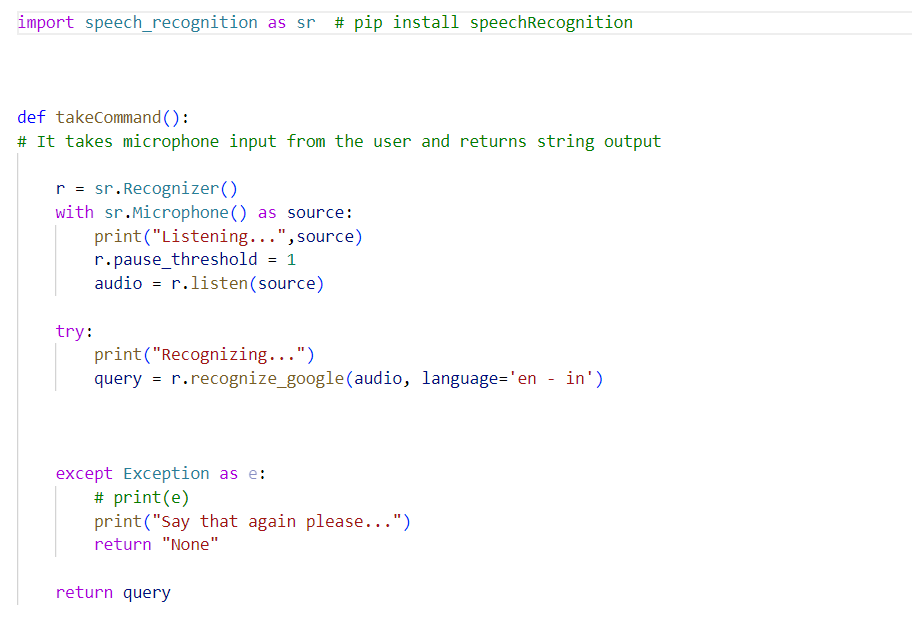


* **Send\_mail()** : This function is used to automatically send the PDF report to the patient via email.

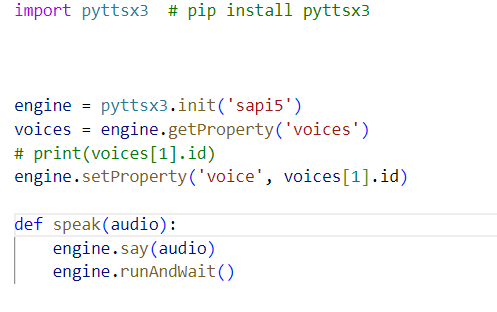
****

1. **Voice Command**

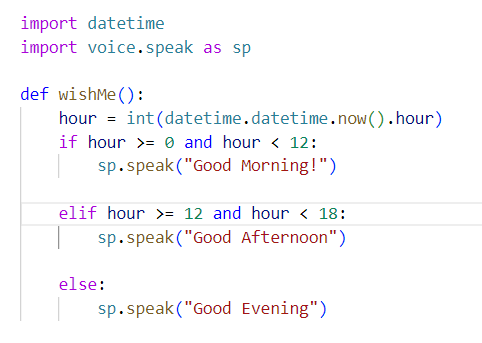
* **Command.py**
* **takeCommand()** : This function is used to take Voice Commands from the user.



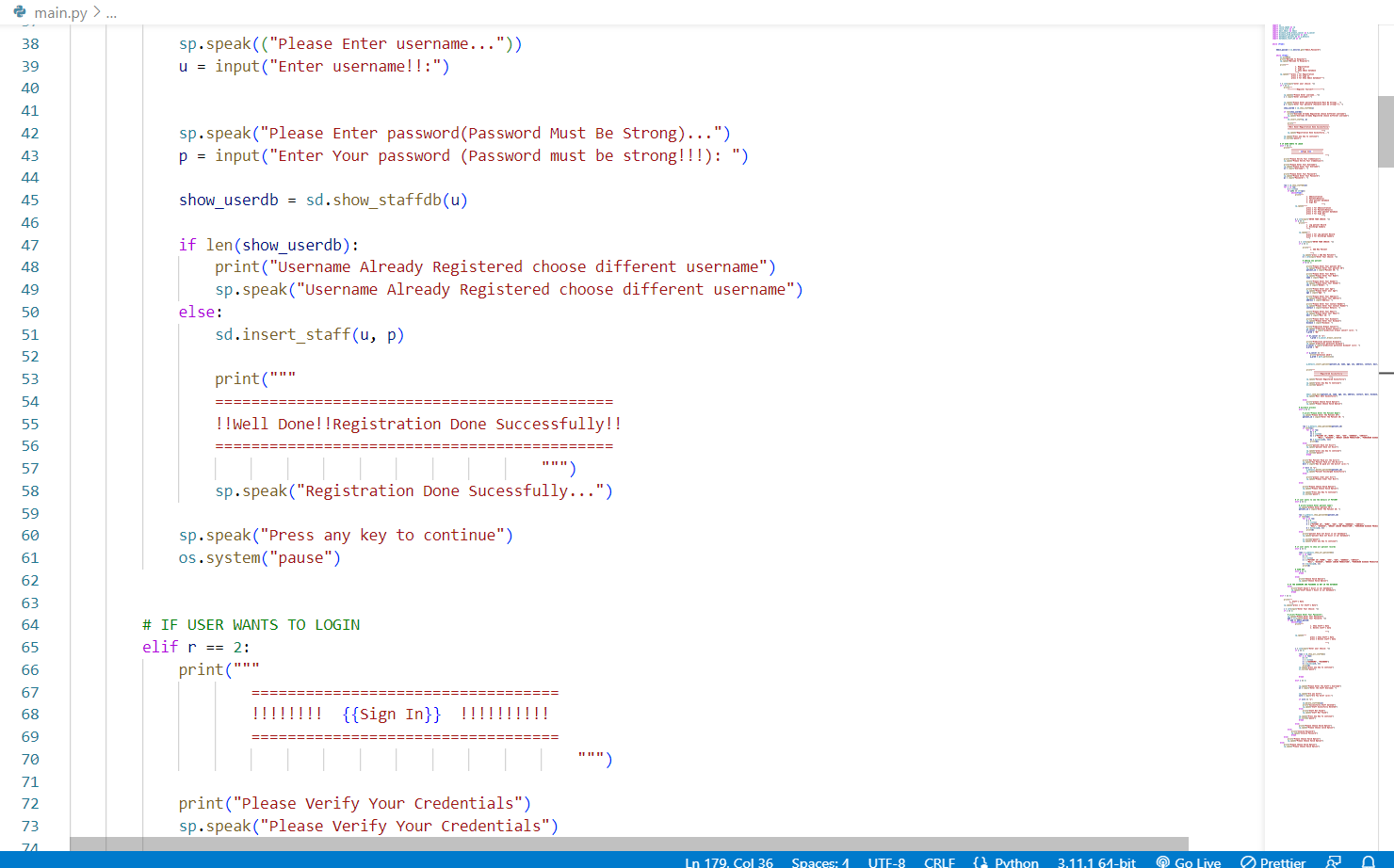
* **Speak.py**
* **Speak()** :

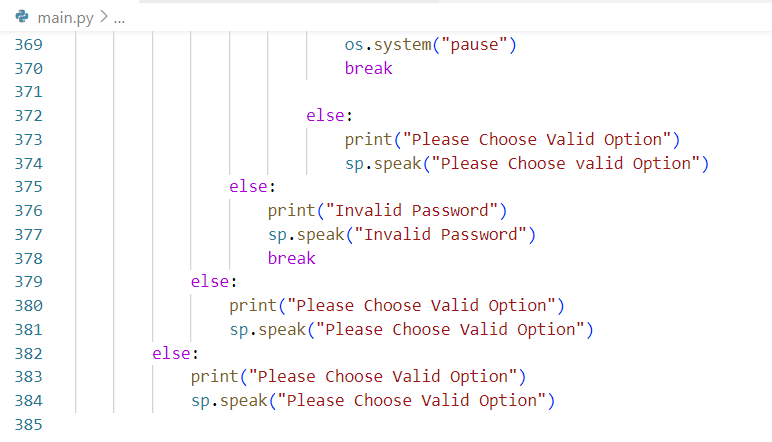
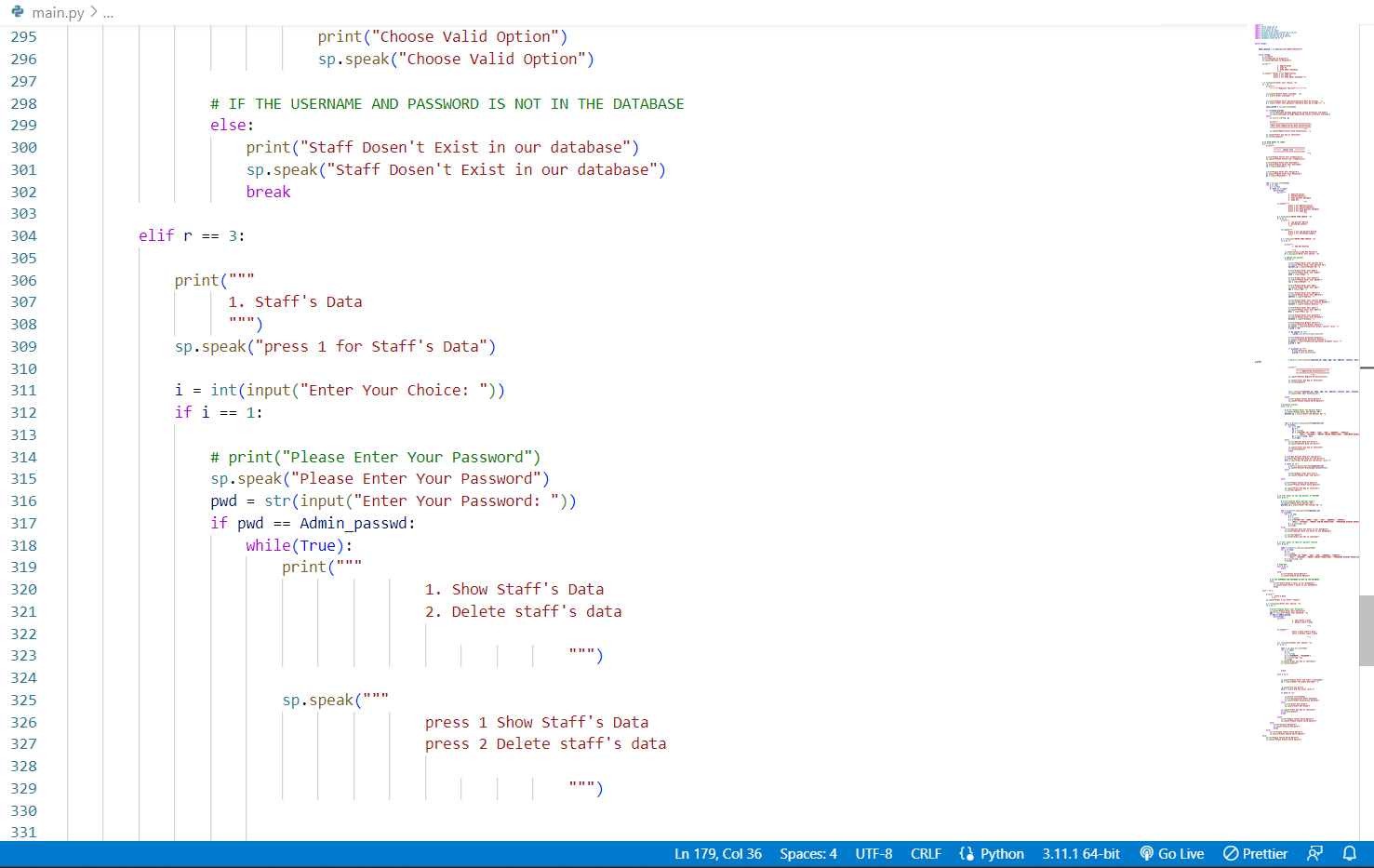
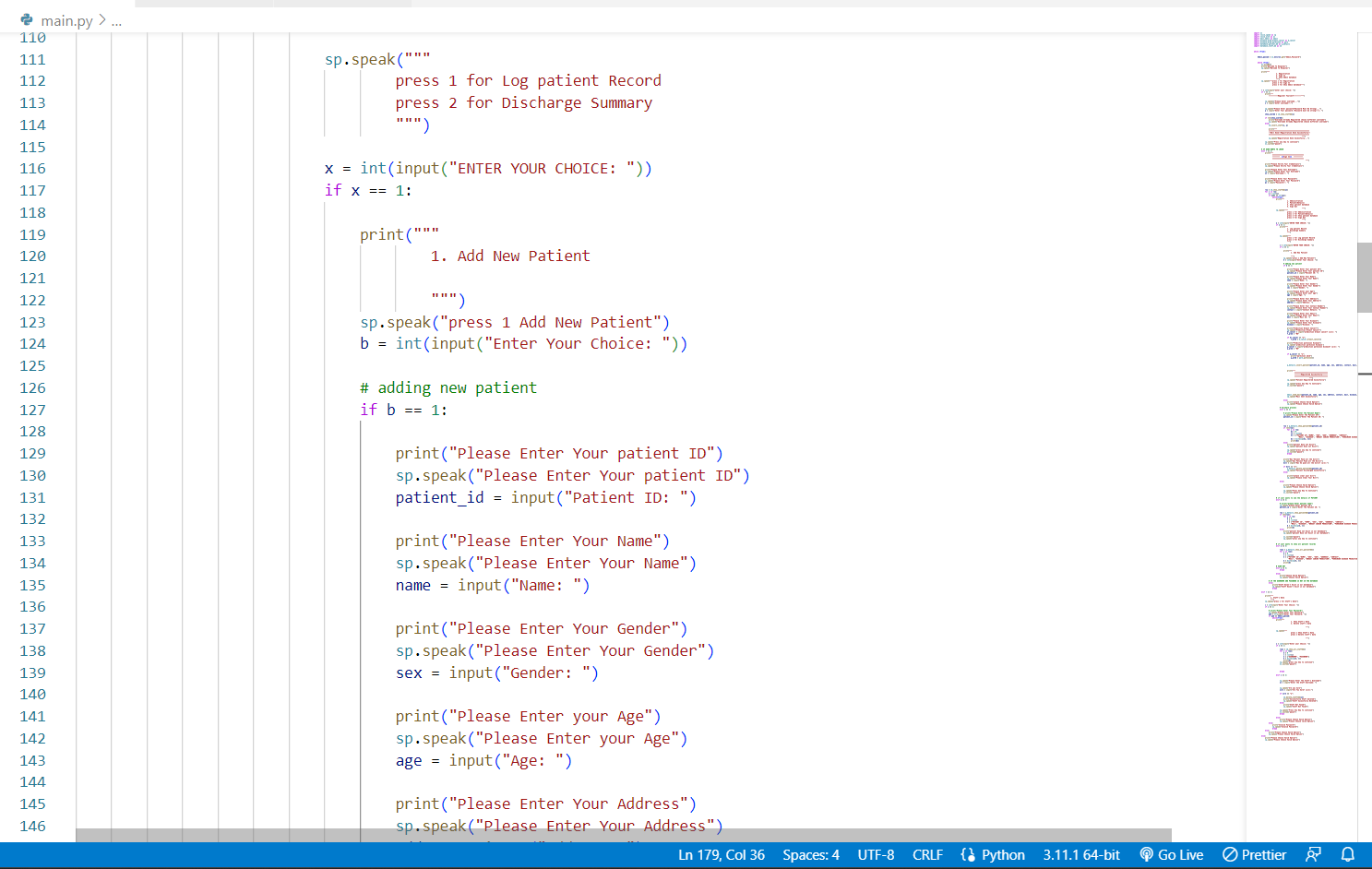
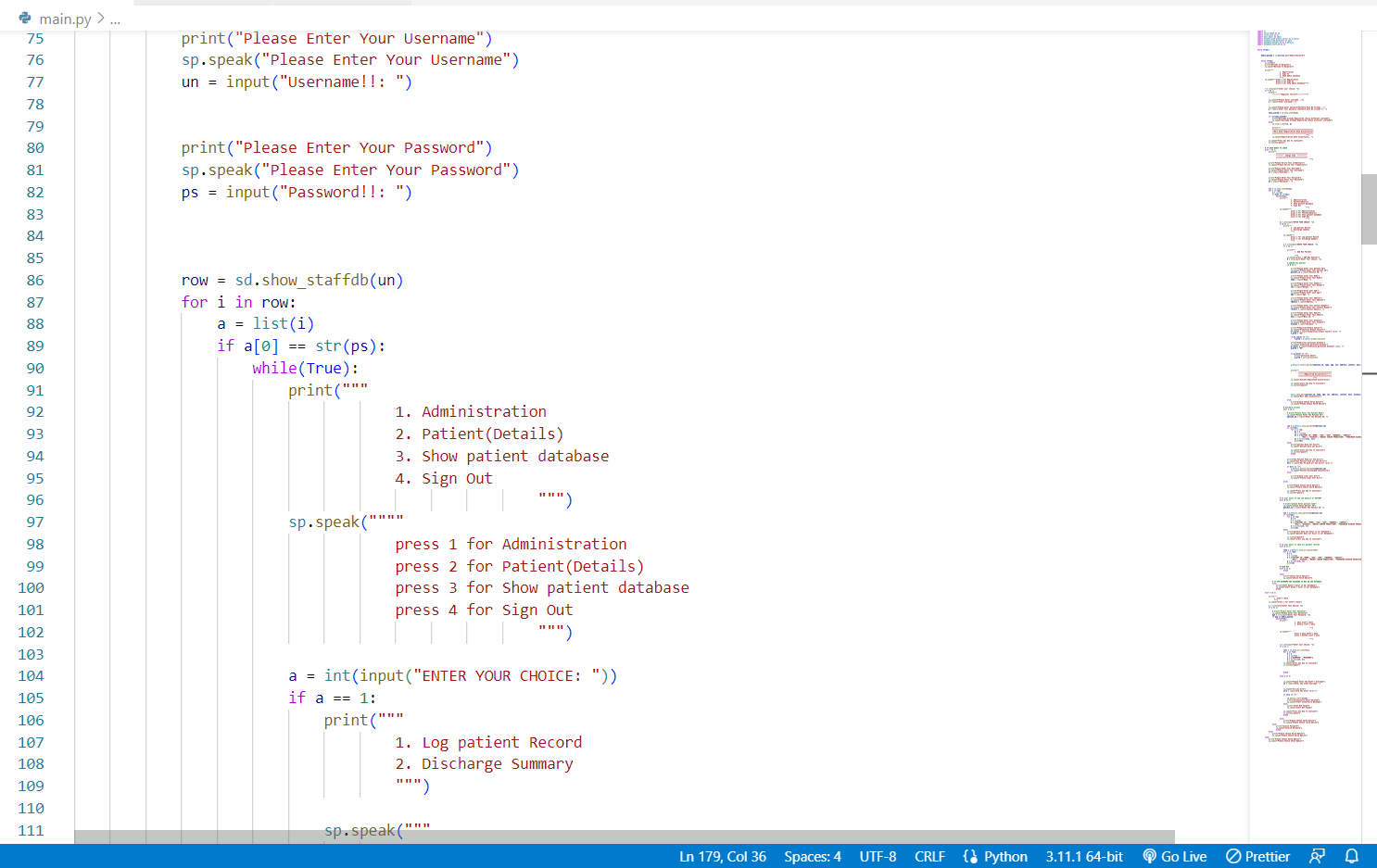
****

* **Wish.py**
* **wishMe()** : This function is used to wish or greet.

****

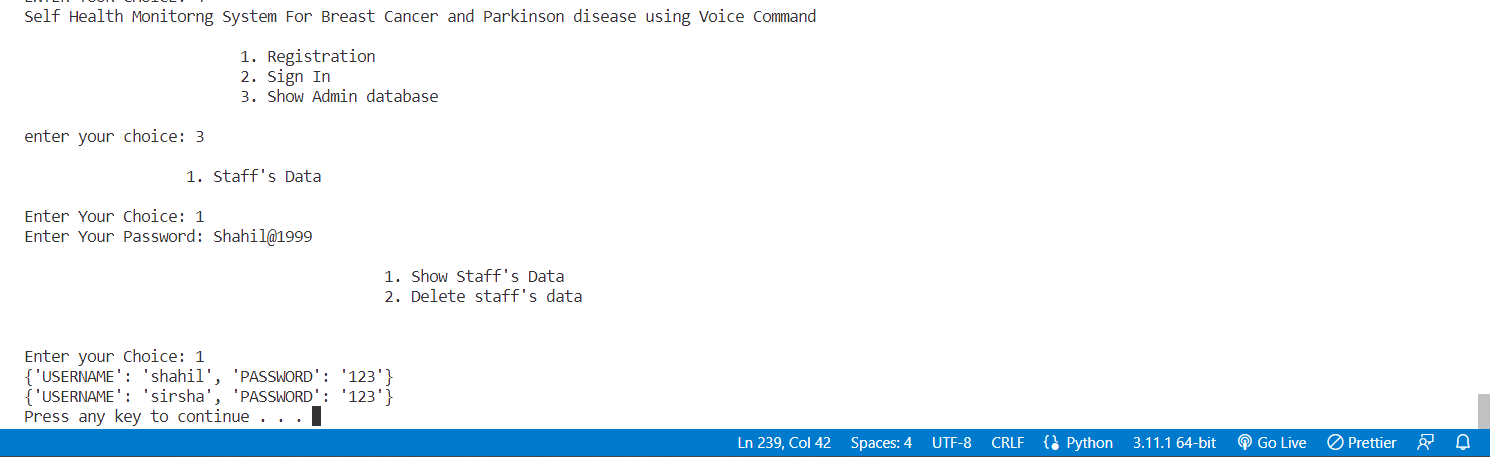
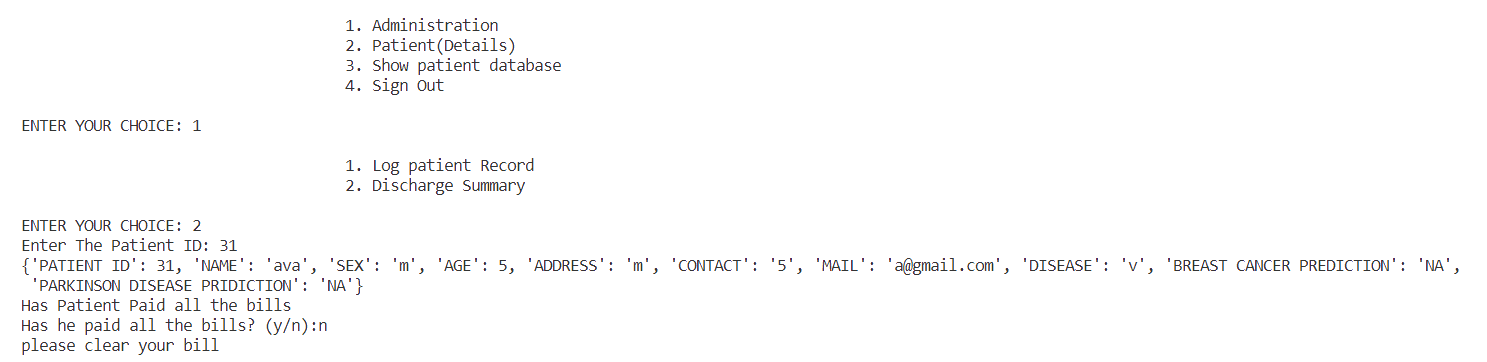
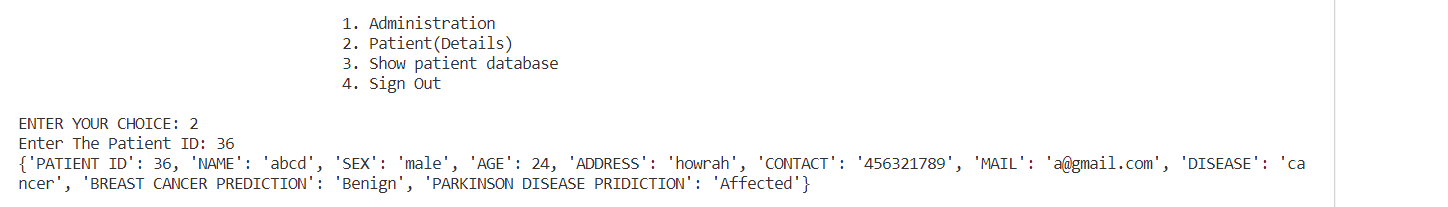
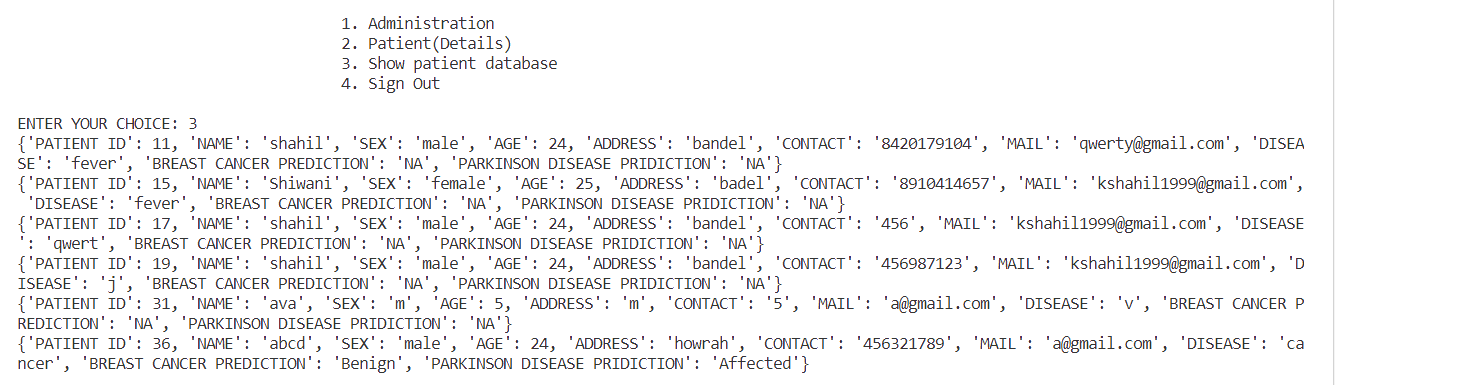
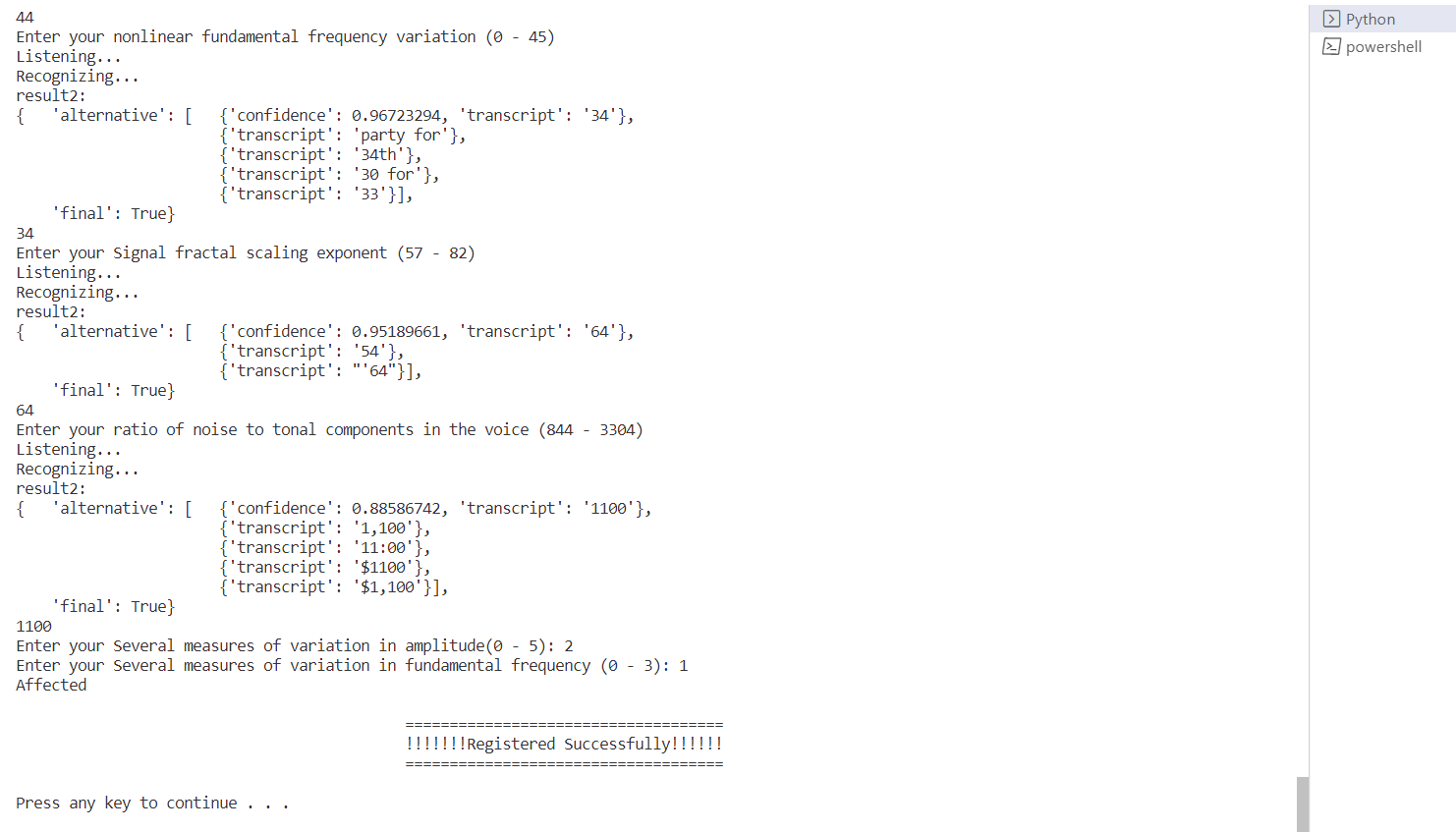
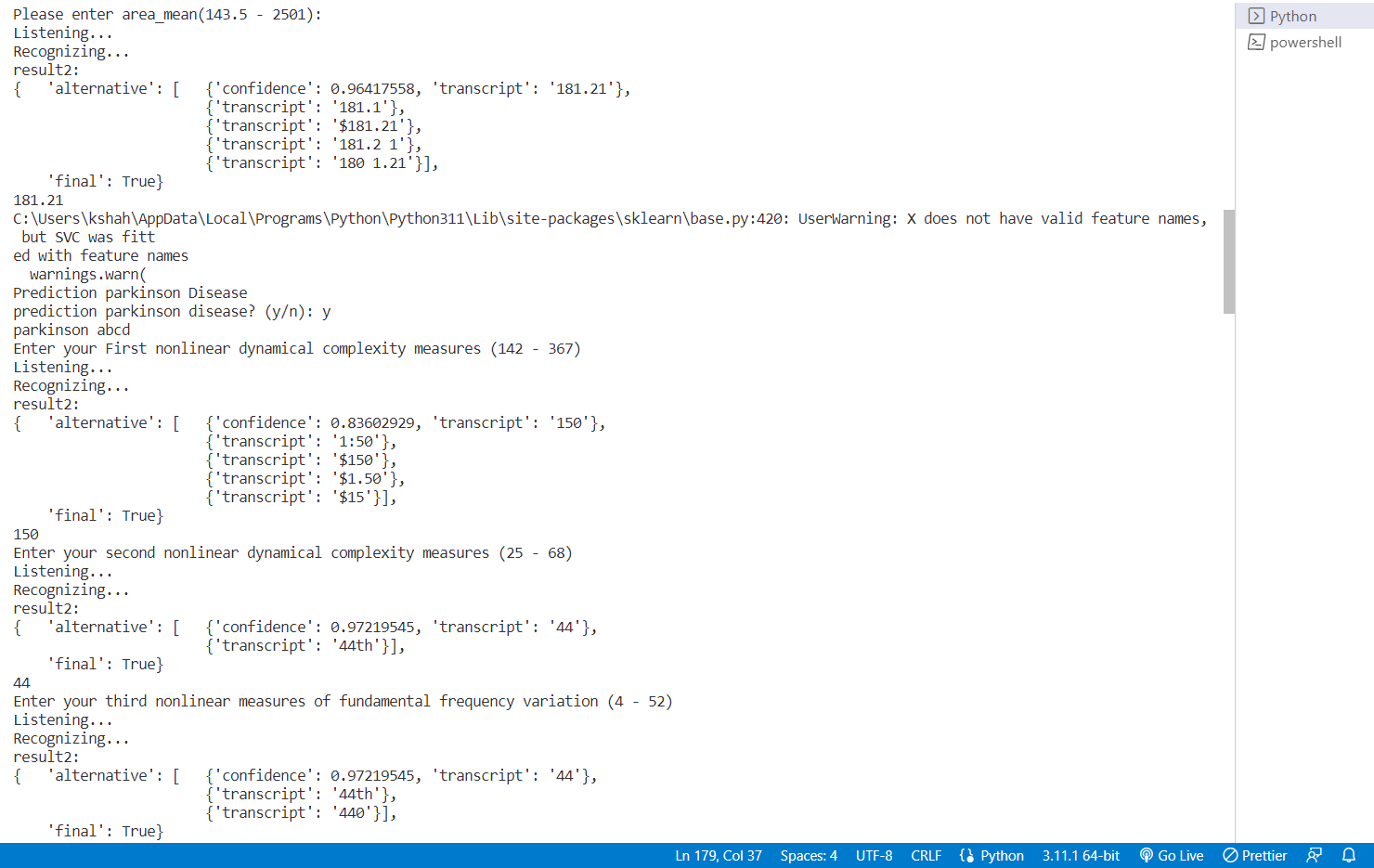
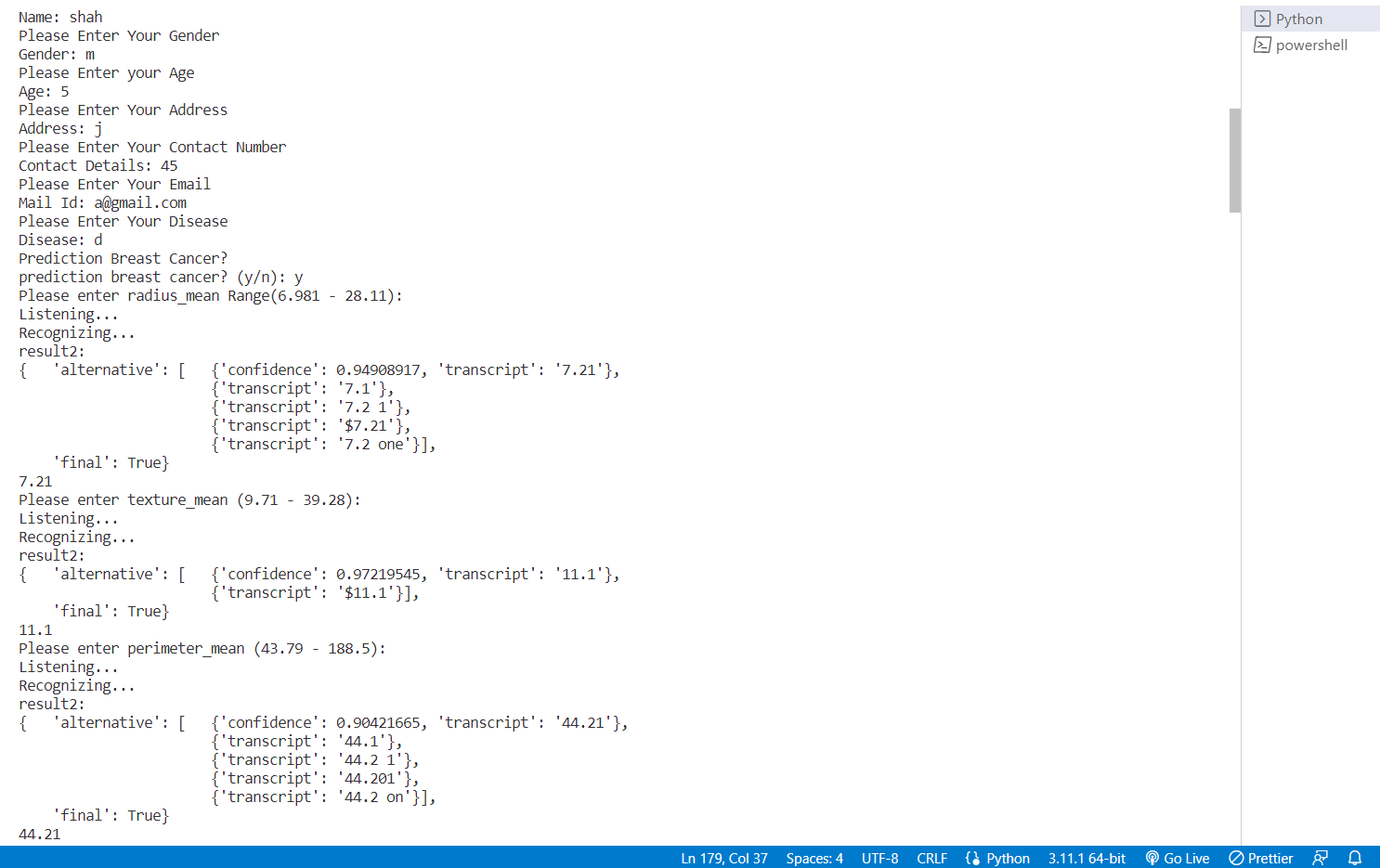
1. **main.py**

****

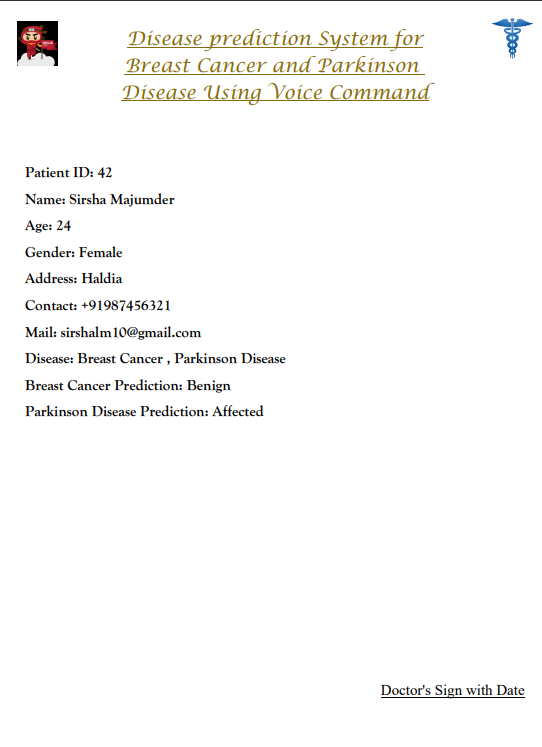
****

**OUTPUTS**





**Health Report**



**EDA OF PARKINSON AND BREAST CANCER DISEASE**

**PARKINSON DISEASE:**

**Features Information:**

Jitter - Several measures of variation in fundamental frequency

Shimmer, - Several measures of variation in amplitude

HNR - measures of ratio of noise to tonal components in the voice

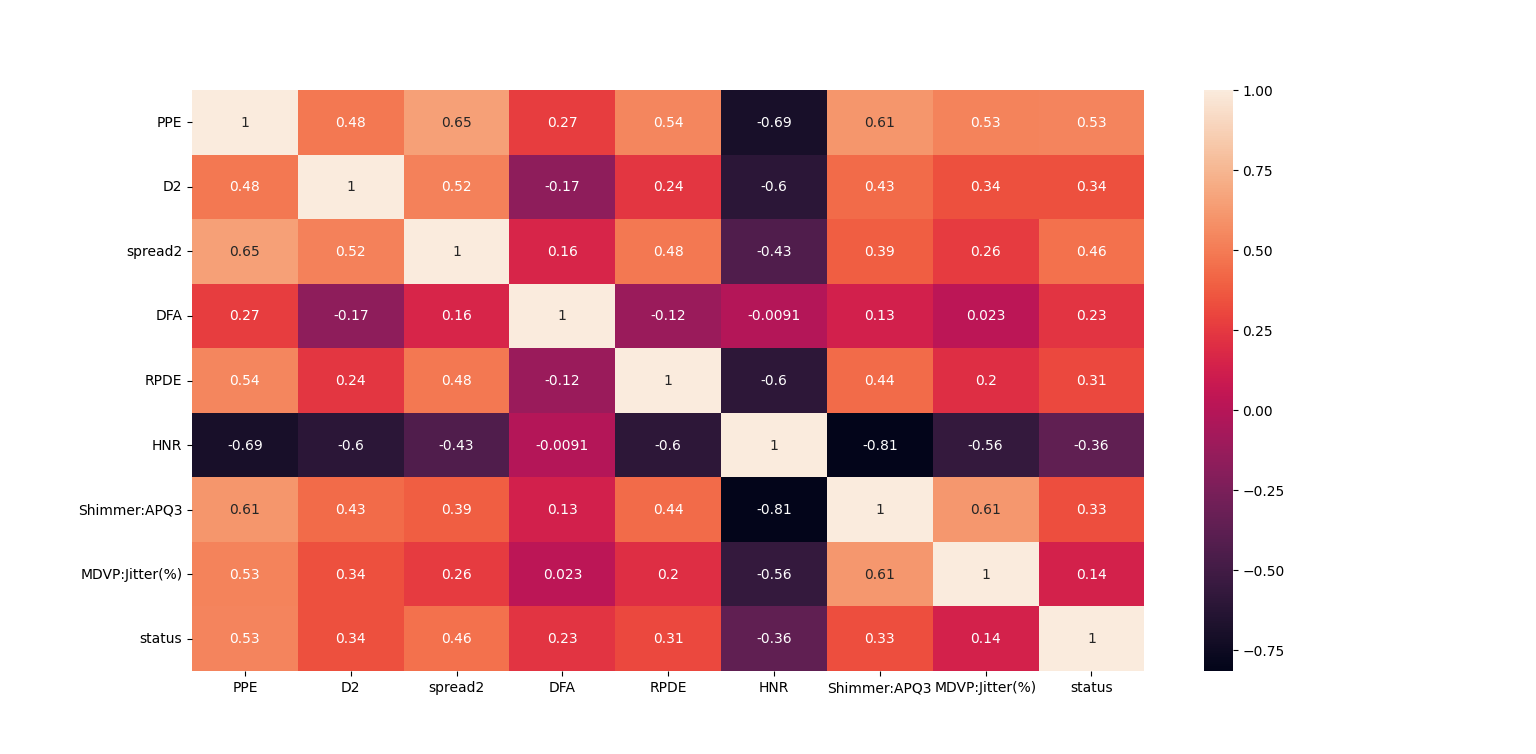
status - Health status of the subject (one) - Parkinson's, (zero) - healthy

RPDE, D2 - Two nonlinear dynamical complexity measures

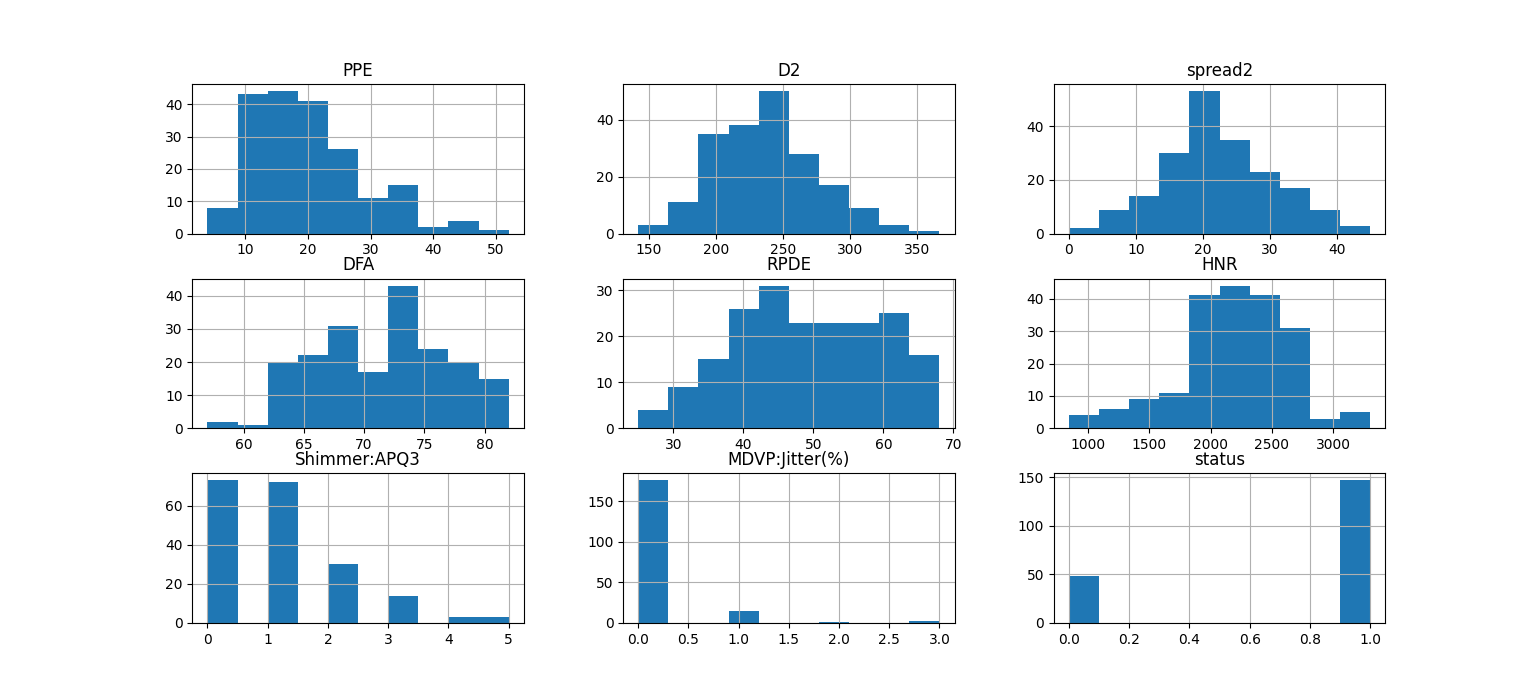
DFA - Signal fractal scaling exponent

spread2, PPE - Two nonlinear measures of fundamental frequency variation

**Heatmap:**

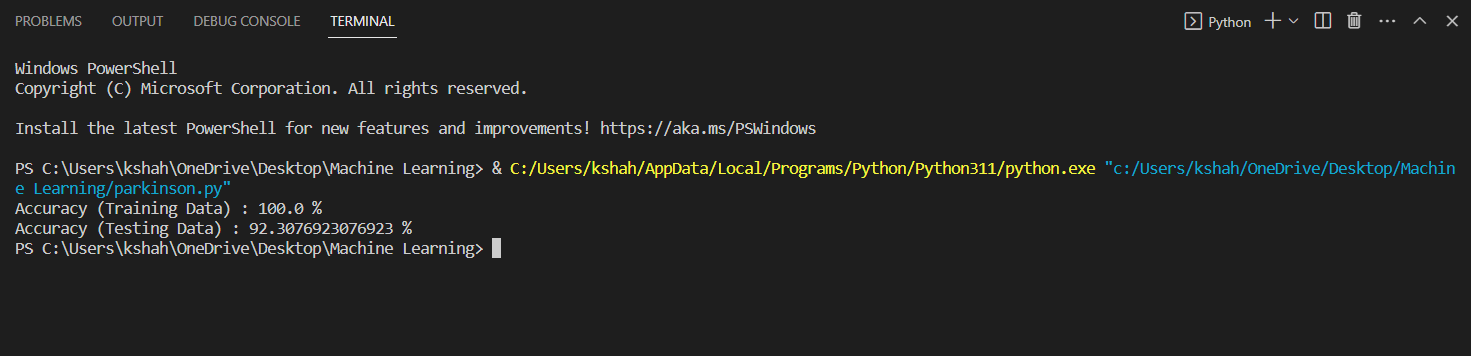
****

**Histogram:**

****

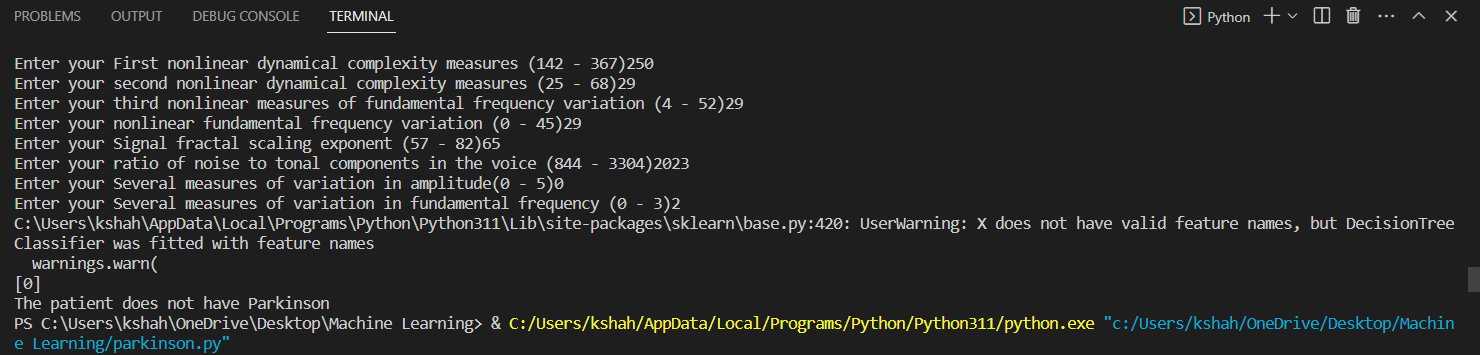
We can see some of the data is normally distributed and most of the attributes are right skewed.

**Accuracy Score:**



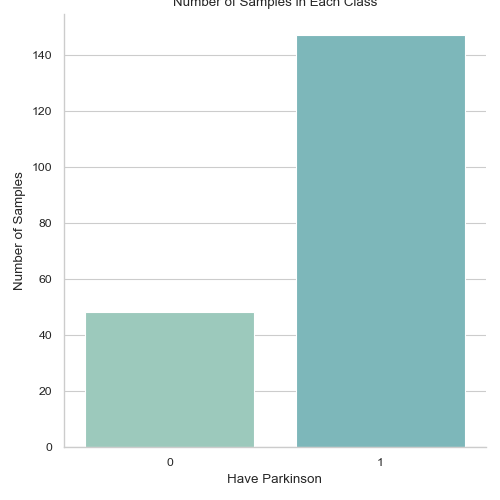
As you can see, we are using DecisionTreeClassifierthe accuracy score in Training Data is 100% and in Testing Data is 92%.

**Output:**

****

As per the input, our algorithm predicted that the patient does not have Parkinson disease.

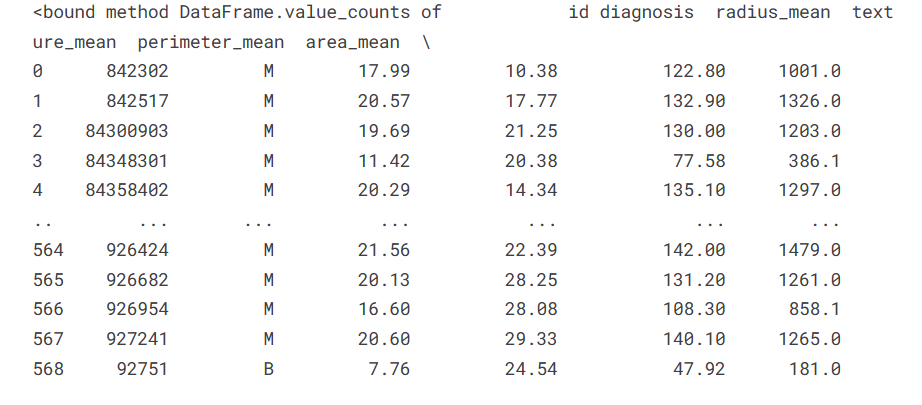
**Bar Diagram:**

****

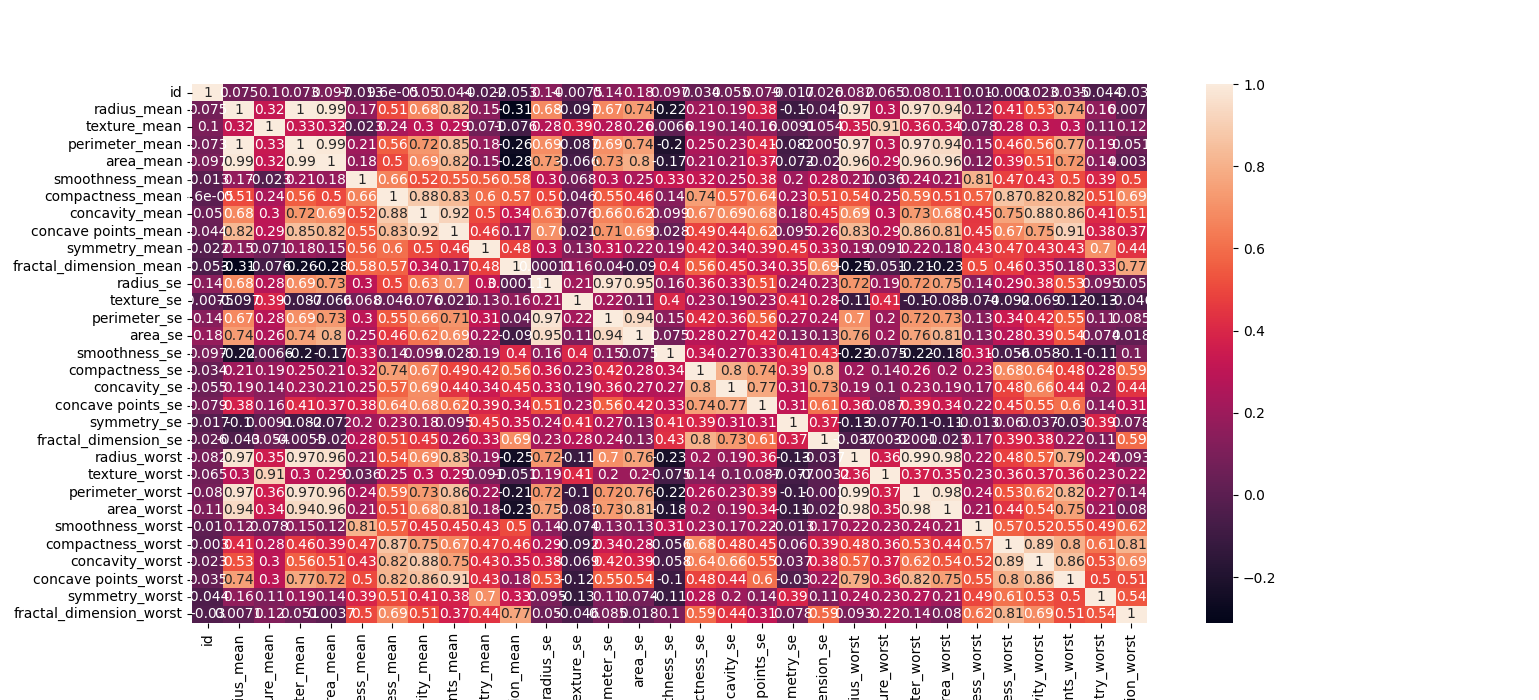
This Bar diagram define number of Output in each class.

**BREAST CANCER:**

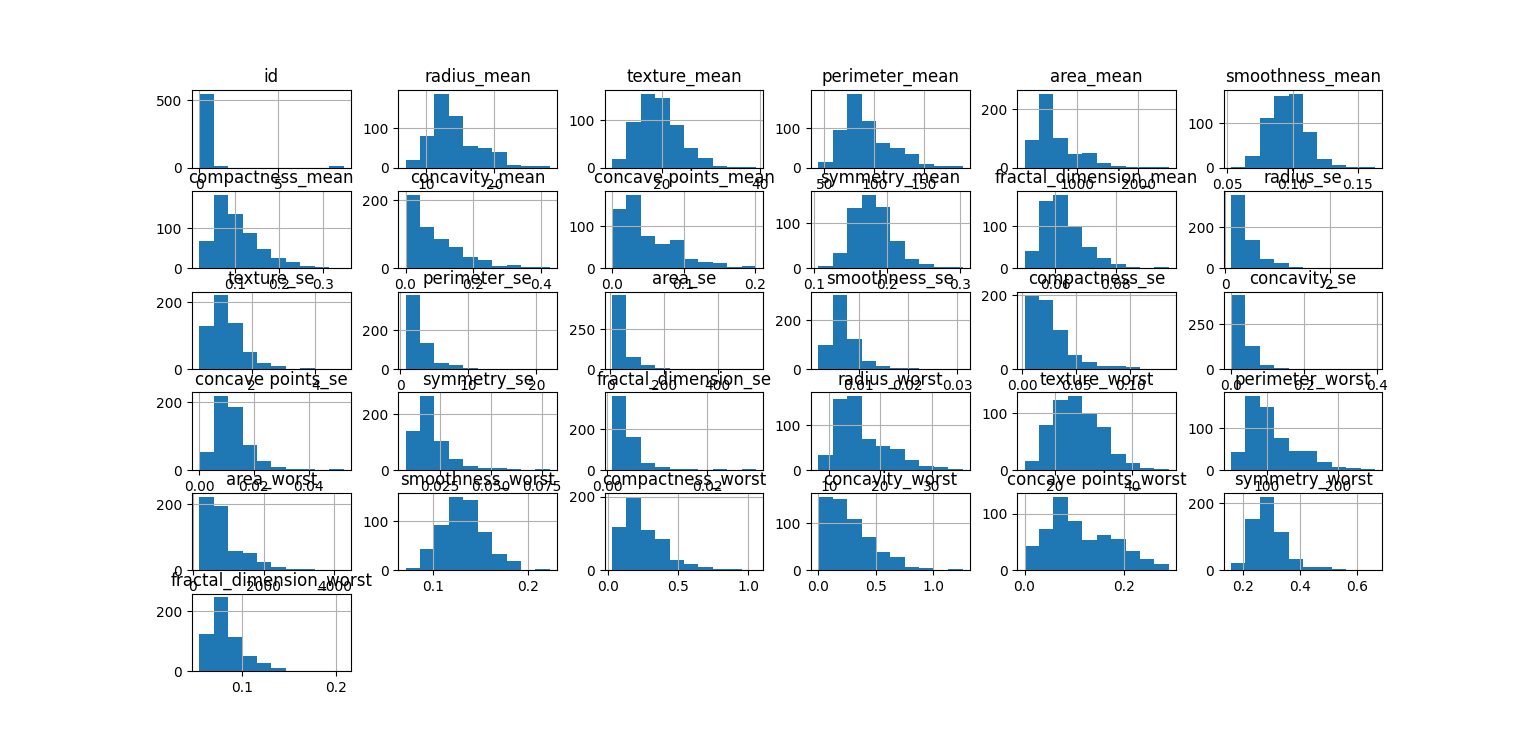
**Data:**

****

**Heatmap:**

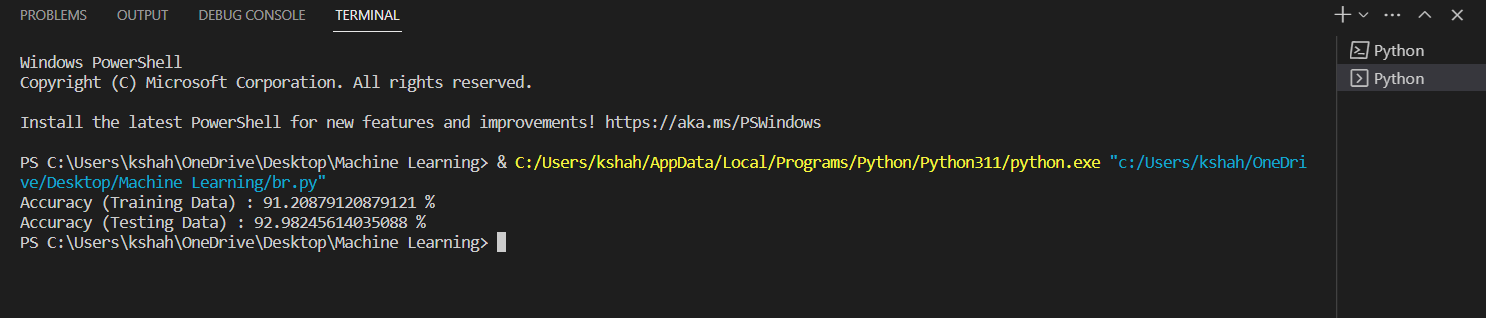
****This is the Heatmap before Feature Engineering.

**Histogram:**

****

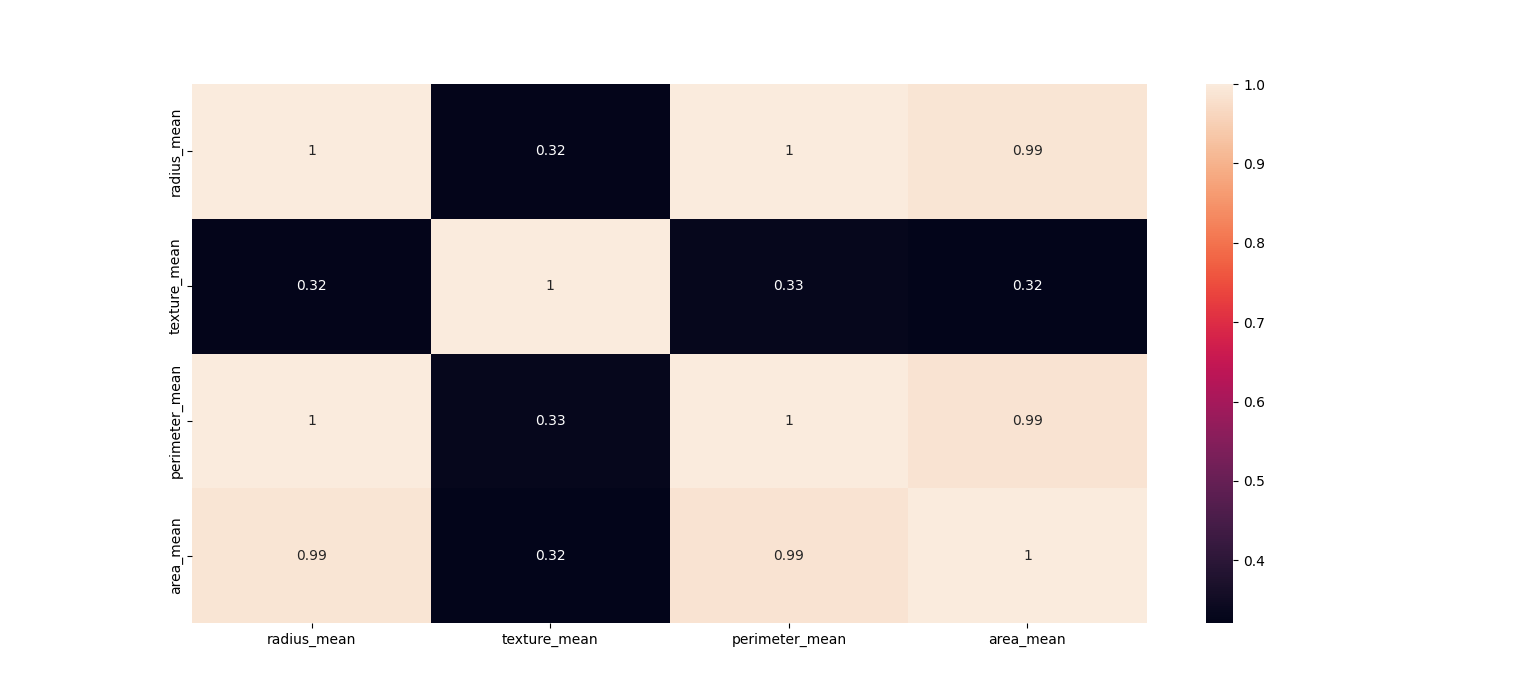
This is the Histogram before Feature Engineering.

**Accuracy Score:**

****

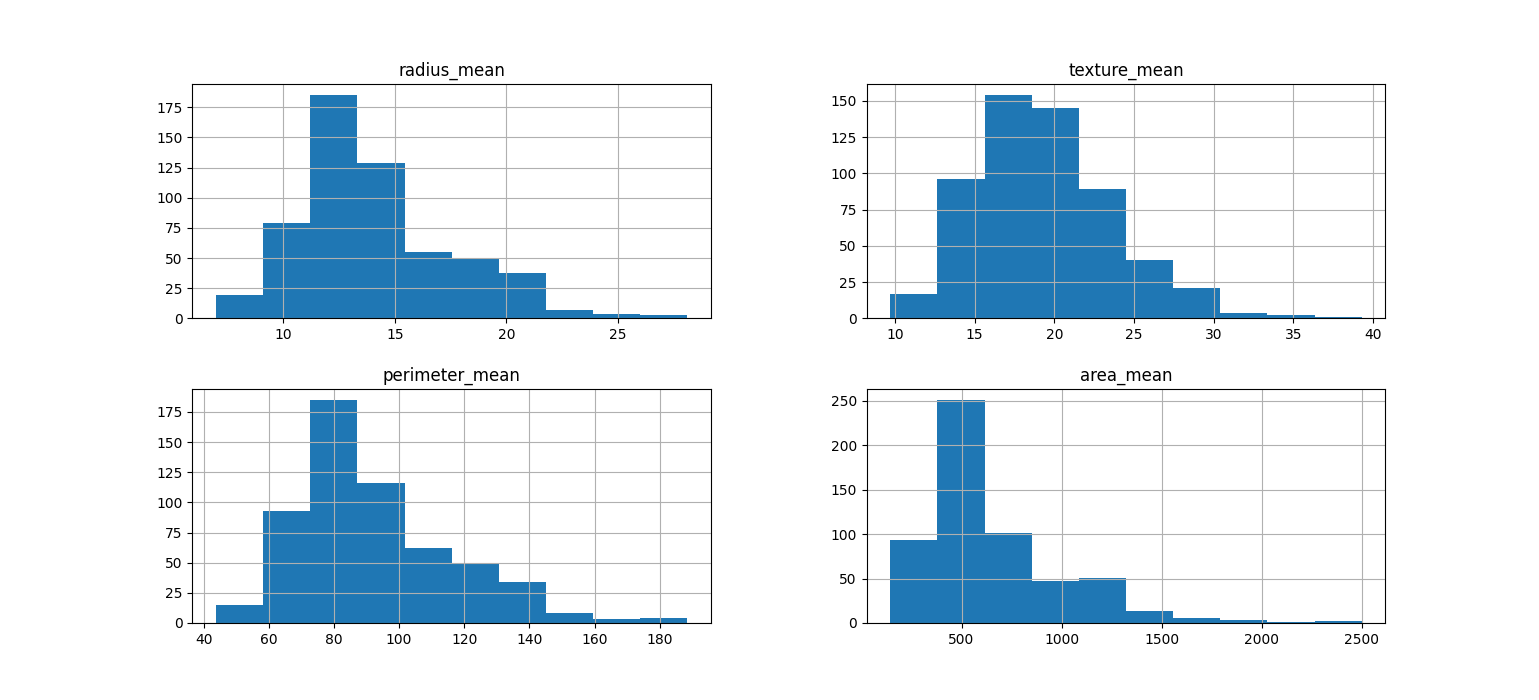
As you can see, we are using Support Vector Machines the accuracy score in Training Data is 100% and in Testing Data is 92%.

**Heatmap:**

****

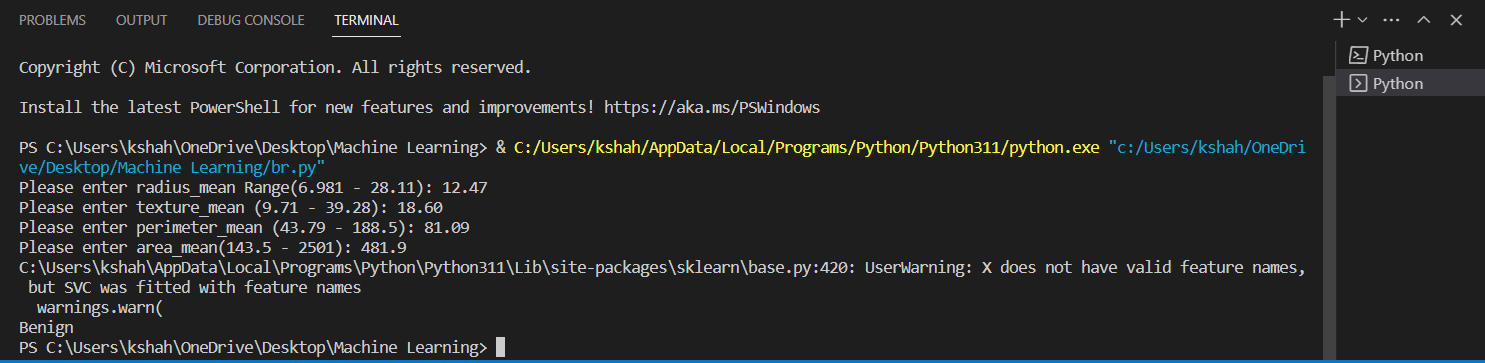
This is the Heatmap after Feature Engineering.

**Histogram:**

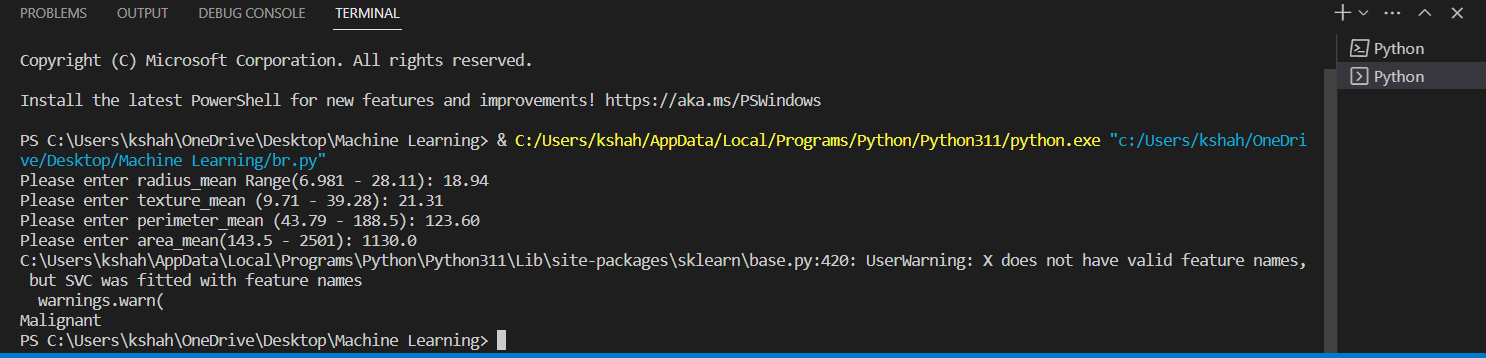
****

This is the after before Feature Engineering.

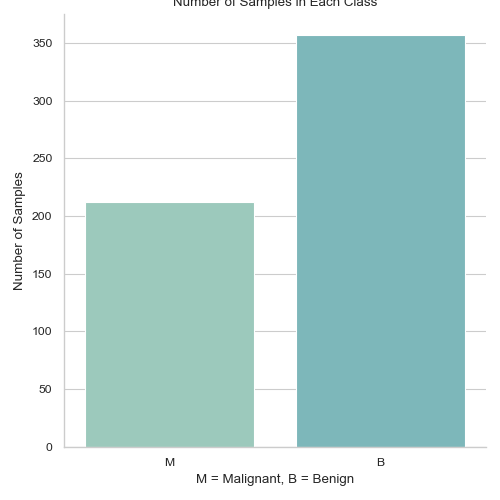
**Output:**

****

As per the input, our algorithm predicted that the patient has Benign type Breast Cancer.

****As per the input, our algorithm predicted that the patient has Malignant Breast Cancer.

**Bar Diagram:**

****

This Bar diagram defines the number and type of Outputs.

**PROPOSED METHODOLOGIES**

**BREAST CANCER:**

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn import svm

import warnings

warnings.filterwarnings("ignore", category=DeprecationWarning)

warnings.filterwarnings("ignore", category=UserWarning)

def breast\_cancer():

    df = pd.read\_csv('C:/Users/kshah/OneDrive/Desktop/test\_major\_project/disease\_pred/br.csv')

 x\_train, x\_test, y\_train, y\_test = train\_test\_split(df.drop(columns=['diagnosis']), df['diagnosis'], test\_size=0.2, random\_state=42)

    clf = svm.SVC(kernel='linear')

    clf.fit(x\_train, y\_train)

    # Accuracy Score on training data

    x\_train\_pred = clf.predict(x\_train)

    training\_data\_accuracy = accuracy\_score(y\_train, x\_train\_pred)

    print('Accuracy (Training Data) :', training\_data\_accuracy\*100,'%')

    # Accuracy Score on test data

    x\_test\_pred = clf.predict(x\_test)

    testing\_data\_accuracy = accuracy\_score(y\_test, x\_test\_pred)

    print('Accuracy (Testing Data) :', testing\_data\_accuracy\*100,'%')

print("Please enter radius\_mean Range(6.981 - 28.11): ")

    sp.speak("Please enter radius\_mean Range(6.981 - 28.11): ")

    radius\_mean = cmd.takeCommand().lower()

    print(radius\_mean)

    sp.speak(radius\_mean)

    print("Please enter texture\_mean (9.71 - 39.28): ")

    sp.speak("Please enter texture\_mean (9.71 - 39.28): ")

    texture\_mean = cmd.takeCommand().lower()

    print(texture\_mean)

    sp.speak(texture\_mean)

 # Splitting the data into testing and training set

    x\_train, x\_test, y\_train, y\_test = train\_test\_split(df.drop(columns=['status']), df['status'], test\_size=0.2, random\_state=42)

    # Data Standardization

    scaler = StandardScaler()

    a = scaler.fit(x\_train)

    x\_train = scaler.transform(x\_train)

    x\_test = scaler.transform(x\_test)

    # Model Training (DecisionTreeClassifier)

    clf = DecisionTreeClassifier()

    clf.fit(x\_train, y\_train)

    print("Please enter perimeter\_mean (43.79 - 188.5): ")

    sp.speak("Please enter perimeter\_mean (43.79 - 188.5): ")

    perimeter\_mean = cmd.takeCommand().lower()

    print(perimeter\_mean)

    sp.speak(perimeter\_mean)

    print("Please enter area\_mean(143.5 - 2501): ")

    sp.speak("Please enter area\_mean(143.5 - 2501): ")

    area\_mean = cmd.takeCommand().lower()

    print(area\_mean)

    sp.speak(area\_mean)

    preds = clf.predict([[radius\_mean, texture\_mean, perimeter\_mean, area\_mean]])

    f\_pred = (' '.join(preds))

    if f\_pred == 'B':

        f\_pred = 'Benign'

    else:

        f\_pred = 'Malignant'

    sp.speak(f\_pred)

    return f\_pred

**PARKINSON DISEASE:**

import os

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

from sklearn.model\_selection import train\_test\_split

from sklearn.preprocessing import StandardScaler

from sklearn import svm

from sklearn.metrics import accuracy\_score

from sklearn.tree import DecisionTreeClassifier

import warnings

warnings.filterwarnings("ignore", category=DeprecationWarning)

warnings.filterwarnings("ignore", category=UserWarning)

def parkinson():

    df = pd.read\_csv('C:/Users/kshah/OneDrive/Desktop/test\_major\_project/disease\_pred/parkinsons.csv')

    # print(df.info())

    # print(df.describe())

    df.isnull().sum()#checking for missing values

    #dropping column axis = 1; dropping row then axis = 0

    #Data Pre-Processing - Seperating Features and Target variables according to their Correlation

    df.drop(["name",'spread1', 'MDVP:Flo(Hz)','MDVP:Fhi(Hz)','MDVP:Fo(Hz)'], axis=1, inplace=True)

    columns = list(df.columns)

    for column in columns:

        if column == "status":

            continue

        filtered\_columns = [column]

        for col in df.columns:

            if (column == col) | (column == "status"):

                continue

            cor\_val = df[column].corr(df[col])

            if cor\_val > 0.75:

                columns.remove(col)

                continue

            else:

                filtered\_columns.append(col)

        df = df[filtered\_columns]

    df.isnull().sum() #checking null value

    # converting Data in the form of hundred

    df.iloc[:,:8] = (df.iloc[:, :8]).mul(100).astype(int)

    # Splitting the data into testing and training set

    x\_train, x\_test, y\_train, y\_test = train\_test\_split(df.drop(columns=['status']), df['status'], test\_size=0.2, random\_state=42)

    # Model Training (DecisionTreeClassifier)

    clf = DecisionTreeClassifier()

    clf.fit(x\_train, y\_train)

 # Model Evaluation

    # Accuracy Score

    # Accuracy Score on training data

    x\_train\_pred = clf.predict(x\_train)

    training\_data\_accuracy = accuracy\_score(y\_train, x\_train\_pred)

    print('Accuracy (Training Data) :', training\_data\_accuracy\*100,'%')

    # Accuracy Score on test data

    x\_test\_pred = clf.predict(x\_test)

    testing\_data\_accuracy = accuracy\_score(y\_test, x\_test\_pred)

    print('Accuracy (Testing Data) :', testing\_data\_accuracy\*100,'%')

print("Enter your First nonlinear dynamical complexity measures (142 - 367)")

    sp.speak("Enter your First nonlinear dynamical complexity measures (142 - 367)")

    D2 = cmd.takeCommand().lower()

    print(D2)

    sp.speak(D2)

    print("Enter your second nonlinear dynamical complexity measures (25 - 68)")

    sp.speak("Enter your second nonlinear dynamical complexity measures (25 - 68)")

    RPDE = cmd.takeCommand().lower()

    print(RPDE)

    sp.speak(RPDE)

    print('Enter your third nonlinear measures of fundamental frequency variation (4 - 52)')

    sp.speak('Enter your third nonlinear measures of fundamental frequency variation (4 - 52)')

    PPE = cmd.takeCommand().lower()

    print(PPE)

    sp.speak(PPE)

    print("Enter your nonlinear fundamental frequency variation (0 - 45)")

    sp.speak("Enter your nonlinear fundamental frequency variation (0 - 45)")

    spread2 = cmd.takeCommand().lower()

    print(spread2)

    sp.speak(spread2)

    print("Enter your Signal fractal scaling exponent (57 - 82)")

    sp.speak("Enter your Signal fractal scaling exponent (57 - 82)")

    DFA = cmd.takeCommand().lower()

    print(DFA)

    sp.speak(DFA)

    print("Enter your ratio of noise to tonal components in the voice (844 - 3304)")

    sp.speak("Enter your ratio of noise to tonal components in the voice (844 - 3304)")

    HNR = cmd.takeCommand().lower()

    print(HNR)

    sp.speak(HNR)

    print("Enter your Several measures of variation in amplitude(0 - 5)")

    sp.speak("Enter your Several measures of variation in amplitude(0 - 5)")

    # Shimar = cmd.takeCommand().lower()

    # print(Shimar)

    # sp.speak(Shimar)

    Shimar = input('Enter variation in amplitude: ')

    print("Enter your Several measures of variation in fundamental frequency (0 - 3)")

    sp.speak("Enter your Several measures of variation in fundamental frequency (0 - 3)")

    # Jitter = cmd.takeCommand().lower()

    # print(Jitter)

    # sp.speak(Jitter)

    Jitter= input('Enter fundamental frequency: ')

    p\_pred = clf.predict([[D2, RPDE, PPE, spread2, DFA, HNR,Shimar, Jitter]])

    predicted = ""

    if p\_pred == 0:

        predicted = 'Not Affected'

    else:

        p\_pred == 1

        predicted = 'Affected'

        return predicted

**DATABASE:**

import mysql.connector

mysql = mysql.connector.connect(host = "XXXXX", user = "XXXXX", passwd = "XXXXX")

mycursor = mysql.cursor()

def insert\_patient(patient\_id, name, sex, age, address, contact, mail, disease, f\_pred, p\_pred):

    mycursor.execute("create database if not exists city\_hospitals")

    mycursor.execute("use city\_hospitals")

     # creating the tables for storing patient details.

    mycursor.execute("create table if not exists patient\_detail(patient\_id int(4) primary key, name varchar(30) ,sex varchar(15),age int(3),address varchar(50),contact varchar(15),mail varchar(40), disease varchar(80), breasr\_cancer\_prediction varchar(20), parkinson\_disease\_prediction varchar(20))")

    # Inserting Patient Details

    mycursor.execute("insert into patient\_detail values('" + patient\_id + "','" + name + "','" + sex + "','" + age + "','" + address + "','" + contact + "','" + mail + "','" + disease + "','" + f\_pred + "','" + p\_pred + "')")

    mysql.commit()

**ACCURACY TABLE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table of accuracy | | | | |
| STUDY DATE | | ACCURACY | SCORE | SOURCE |
| May  2019 | | Train | 83.44% | https://www.kaggle.com/code/parhamzm/parkinson-s-disease-pd-classification/notebook#notebook-container |
| Test | 85.53% |
| Feb  2023 | | Train | 84.65% | https://www.kaggle.com/code/akanksha10/detection-of-parkinson-s-disease |
| Test | 86.8% |
| Jan  2021 | | Train | 93.57% | https://www.kaggle.com/code/vikasukani/detecting-parkinson-s-disease-machine-learning |
| Test | 96.66% |
| July  2022 | | Train | 88.46% | https://www.youtube.com/watch?v=ys\_mVbkaokE |
| Test | 87.17% |
| June  2016 | | Train | 74.60% | https://www.kaggle.com/code/lykin22/parkinson-s-disease-based-on-voice-recording#Parkinson's-disease-based-on-voice-recording |
| Test | 74.55% |
| Feb  2015 | | Train | 95.38% | https://www.researchgate.net/profile/Anil-Kumar-544/google scholar |
| Test | 94.72% |
| Feb  2018 | | Train | 95.38% | https://www.researchgate.net/profile/Anil-Kumar-544/google scholar |
| Test | 94.72% |
| Aug  2017 | | Train | 97% | https://www.youtube.com/watch?v=eKy3KgRgDkQ |
| Test | 98% |
| Sep  2020 | | Train | 88.46% | https://www.youtube.com/watch?v=CQLkX4utdIU |
|  | Test | | 87.11% |
| April, 2022 | Train | | 88.46% | https://github.com/akashdeep364/Parkinson-s-Disease-Detection |
| Test | | 87.17% |

**APPLICATIONS**

This project is based on a trending technology of the present times and has many applications –

1. One of the most important applications of this project is that, the accessibility to the Hospital Staff and Patient records are available in Database can easily accessible making it more user friendlyhelping the hospital administration to manage data even during rush hours smoothly.

2. This project paves the path for a smooth guidance to all the health-conscious individuals specially differently abled people irrespective of their age and health conditions.

3. The users are also given relief from the hassle of storage issues when it comes to using this application as it provides the opportunity of online data storage.

4. Data can be inserted, updated when required, deleted, and can also be saved separately in Database tables uncomplicated distinguish between information of users.

5.Fast and Early Prediction of Life taking Diseases Like Breast Cancer and Parkinson Diseases.

6.Automated Mail Sending Feature makes it more reliable and time saving.

**STRENGTHS**

* This Project is completely based on Human Voice Command.Because of it’s Voice Controlled feature it is more beneficial for People with disability.
* Easy to Predict the onset of Parkinson disease and Breast Cancer.
* Uses Machine Learning Algorithm for the diseases Prediction and gives fast results and this can be used for Prediction at a rare critical situation like unavailability of a Neurologist and Oncologist.
* User can send the Reports to the Patients over mail. Reduces Paperwork & acts as a Document that can be accessed from any device at any time.
* Easy to access any Staff’s or Patient’s data from anywhere in the world via Authorised Login.

**LIMITATIONS**

* The high cost of software development and deployment.
* Complex Machine Learning Algorithm in terms of User Experience.
* Fear of data security breach.
* Difficulty in migrating from manual processes, because both staff and patients are used to the manual processes and so are unable to speedily cope with the new system.
* Lack of IT-friendly medical personnel is also presenting several challenges.
* Sometimes Predicated result may not gives 100% accuracy.
* Needs Internet connectivity to send the reports over mail.

**CONCLUSION**

With our proposed system, comparatively a good and higher accuracy is achieved. This is then used by researchers, physicians, hospitals, healthcare centres or doctors in order to provide the best treatment and medical care for the patients. Hence machine learning when used in healthcare can lead to an effective treatment and the patient is also well taken care of. Here we try to implement some of the functions of machine learning in healthcare into our system. In place of direct diagnosis, when a patient’s disease is to be predicted then machine learning is implemented using certain algorithms. In this way, healthcare can be made much better and advanced. When we compare the different algorithms used for disease prediction from our dataset and the output we expect we get the best accuracy with Decision Tree Classifier and Support Vector Machine(SVM) whereas LDA algorithm had the lowest performance when compared to the other algorithms. Machine Learning (ML) gives us different methods and techniques that can make the issue of diagnostic problems easy and simple by modernizing different medical domains. Today, ML is used largely to predict, analyze clinical works and process data analysis like error detection in the dataset and for dealing with incorrect data present in our system. It is no doubt that implementing ML algorithm helped integrating computer system in the industry of healthcare to facilitate and enhance the work of doctors and finally leading to improve the efficiency level and quality of our medical care for the respective patients.

**FUTURE DEVELOPMENT**

In near Future, we are thinking to develop a Single Page User Management Website for controlling and managing all the Frontend flows which will enhance User Interface and User Experience. We will also deploy our Project on AWS cloud so that it can easily be accessible independent of Single User Machine. We will add more number of Diseases prediction algorithm to make it more useful.Can be developed as a self - disease prediction system so that early stages are recognized faster decreasing the immortality rate or health risks.

In today’s world most of the data is computerized, the data is distributed, and it is not utilizingproperly. With the help of the already present data and analysing it, we can also use for un-known patterns. The primary motive of this project is the prediction of diseases with high rate of accuracy. For predicting the disease, we can use logistic regression algorithm, naive Bayes,sklearn in machine learning. The future scope of the paper is the prediction of diseases by usingadvanced techniques and algorithms in less time complexity.A technology called CAD is more beneficial as sometimes systems are better diagnosticsthan Doctors. Machine Learning and its different branches are used in Cancer detection as well.It helps or can say assist in making decisions on critical cases or on therapies. Artificial intelligence plays an important role in development of many health related procedure or methods.Artificial intelligence is very common now a days in surgeries, like Robotics surgery. Since were in the circumstances of growing population, we must need technology which can help usto meet the expectations of the patients, their flawless cure, their better health and their smooth and easy approachable access to healthcare industries to heal and get well soon!

**REFERENCES**

The Project Synopsis has been prepared by taking help from –

* A.Davis, D., V.Chawla, N., Blumm, N., Christakis, N., & Barbasi, A. L. (2008). Predicting Individual Disease Risk Based On Medical History.
* Adam, S., & Parveen, A. (2012). Prediction System For Heart Disease Using Naive Bayes. Al-Aidaroos, K., Bakar, A., & Othman, Z. (2012). Medical Data Classification With Naive Bayes
* Approach. Information Technology Journal .
* Darcy A. Davis, N. V.-L. (2008). Predicting Individual Disease Risk Based On Medical History.
* JyotiSoni, Ansari, U., Sharma, D., & Soni, S. (2011). Predictive Data Mining for Medical Diagnosis: An Overview Of Heart Disease Prediction.
* K.M. Al-Aidaroos, A. B. (n.d.).
* K.M. Al-Aidaroos, A. B. (n.d.). 2012. Medical Data Classification With Naive Bayes Approach .
* NisharBanu, MA; Gomathy, B;. (2013). Disease Predicting System Using Data Mining Techniques.
* Youtube Link: <https://www.youtube.com/playlist?list=PLKnIA16_Rmvbr7zKYQuBfsVkjoLcJgxHH>
* Wikipedia.
* Our Project Guide

# **INDIVIDUAL CONTRIBUTION TO THE PROJECT**

|  |  |
| --- | --- |
| **NAME OF THE PROJECT MEMBER** | **CONTRIBUTION** |
| **MUJAHID ALI ANSARI** | **SQL, MACHINE LEARNING** |
| **ASHUTOSH KUMAR YADAV** | **DOCUMENTATION** |
| **SHAHIL KUMAR CHOURASIA** | **PYTHON, MACHINE LEARNING** |
| **SONU ROUTH** | **TESTING, DBMS** |
| **SIRSHA MAJUMDER** | **PYTHON,**  **DOCUMENTATION** |

**THANK YOU**