PROJECT
REPORT

1. **INTRODUCTION**
   a. Project Overview
   b. Purpose
2. **LITERATURE SURVEY**
   a. Existing problem
   b. References
   c. Problem Statement Definition
3. **IDEATION & PROPOSED SOLUTION**
   a. Empathy Map Canvas
   b. Ideation & Brainstorming
   c. Proposed Solution
   d. Problem Solution fit
4. **REQUIREMENT ANALYSIS**
   a. Functional requirement
   b. Non-Functional requirements
5. **PROJECT DESIGN**
   a. Data Flow Diagrams
   b. Solution & Technical Architecture
   c. User Stories
6. **PROJECT PLANNING & SCHEDULING**
   a. Sprint Planning & Estimation
   b. Sprint Delivery Schedule
   c. Reports from JIRA
7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**
   a. Feature 1
   b. Feature 2
   c. Database Schema (if Applicable)
8. **TESTING**
   a. Test Cases
   b. User Acceptance Testing
9. **RESULTS**
   a. Performance Metrics
10. **ADVANTAGES & DISADVANTAGES**
11. **CONCLUSION**
12. **FUTURE SCOPE**
13. **APPENDIX**

Source Code

GitHub & Project Demo Link

1.INTRODUCTION

a.Project Overview

Parkinson's disease is a brain disorder that causes unintended or uncontrollable movement, such a shaking, stiffness, and difficulty with balance and coordination. Symptoms usually begin gradually and worsen over time. As the disease progresses, people may have difficulty walking and talking. The recent report of the World Health Organization shows a visible increase in the number and health burden of Parkinson's disease patients increases rapidly. In China, this disease is spreading so fast and estimated that it reaches half of the population in the next 10 years. Classification algorithms are mainly used in the medical field for classifying data into different categories according to the number of characteristics. Parkinson's disease is the second most dangerous neurological disorder that can lead to shaking, shivering, stiffness, and difficulty walking and balance. It caused mainly due by the breaking down of cells in the nervous system. Parkinson's can have both motor and non-motor symptoms. The motor symptoms include slowness of movement, rigidity, balance problems, and tremors. If this disease continues, the patients may have difficulty walking and talking. The non-motor symptoms include anxiety, breathing problems, depression, loss of smell, and change in speech. If the above-mentioned symptoms are present in the person then the details are stored in the records. In this paper, the author considers the speech features of the patient, and this data is used for predicting whether the patient has Parkinson's disease or not. Neurodegenerative disorders are the results of progressive tearing and neuron loss in different areas of the nervous system. Neurons are functional units of the brain. They are contiguous rather than continuous. A good healthy looking neuron as shown in fig 1 has extensions called dendrites or axons, a cell body, and a nucleus that contains our DNA. DNA is our genome and a hundred billion neurons contain our entire genome which is packaged into it. When a neuron gets sick, it loses its extension and hence its ability to communicate which is not good for it and its metabolism becomes low so it starts to accumulate junk and it tries to contain the junk in the little packages in little pockets. When things become worse and if the neuron is a cell culture it completely loses its extension, becomes round and full of vacuoles. This work deals with the prediction of Parkinson's disorder which is now a day is tremendously increasing incurable disease. Parkinson's disease is a most spreading disease which gets its name from James Parkinson who earlier described it as a paralysis agitans and later gave his surname was known as PD. It generally affects the neurons which are responsible

for overall body movements. The main chemicals are dopamine and acetylcholine which affect the human brain. There is a various environmental factor which has been implicated in PD below are the listed factor which caused Parkinson's disease in an individual. · Environmental factors: Environment is defined as the surroundings or the place in which an individual lives. So the environment is the major factor that will not only affects the human's brain but also affects all the living organism who lives in the vicinity of it. Many types of research and evidence have proved that the environment has a big hand in the development of neurodegenerative disorders mainly Alzheimer's and Parkinson's. There are certain environmental factors that are influencing neurodegenerative disorder with high pace are:- · Exposure to heavy metals (like lead and aluminum) and pesticides. · Air Quality: Pollution results in respiratory diseases. · Water quality: Biotic and Abiotic contaminants present in water lead to water pollution. · Unhealthy lifestyle: It leads to obesity and a sedentary lifestyle. · Psychological stress: It increases the level of stress hormone that depletes the functions of neurons. · Brain injuries or Biochemical Factors: The brain is the control center of our complete body. Due to certain trauma, people have brain injuries which leads some biochemical enzymes to come into the picture which provides neurons stability and provides support to some chromosomes and genes in maintenance. · Aging Factor: Aging is one of the reasons for the development of Parkinson's disease. According to the author in India, 11,747,102 people out of 1, 065, 070, 6072 are affected by Parkinson's disease. · Genetic factors: Genetic factor is considered as the main molecular physiological cause which leads to neurodegenerative disorders. The size, depth, and effect of actions of different genes define the status or level of neurodegenerative disease which increases itself gradually over time. Mainly the genetic factors which lead to Neurodegenerative disorders are categorized into pharmacodynamics and pharmacokinetics. · Speech Articulation factors: Due to the condition associated with Parkinson's disease (rigidity and bradykinesia), some speech-language pathology such as voice, articulation and swallowing alterations are found. There are various ways in which Parkinson's disease (PD) might affect the individual. · The voice get breathy and softer. · Speech may be smeared. · The person finds difficulty in finding the right words due to which speech becomes slower.

b.Purpose

The key purpose of the Personalized Parkinson Project is **to contribute to the understanding of the differences in etiology, pathophysiology, phenotypic diversity, and disease progression among individual PD patients**. Although previous cohort studies have largely contributed to elucidate the differences between PD patients and healthy control subjects [8, 45,46,47], we remain unable to understand the basis for the large diversity of phenotypes and variability in progression rates among PD patients. It remains unclear why some PD patients stay functional and independent long into the disease, while others progress to significant motor and/or cognitive impairment and are unable to live unassisted relatively early in the course of the disease. The 2-year follow-up captures inter-individual differences in the speed of disease progression on numerous clinically relevant outcomes, including motor- and non-motor symptoms and endpoints that are relevant in the earlier phases of the disease, such as the ability to work. The PPP has several unique elements: an unbiased approach to patient selection, with purposely broad inclusion criteria (also allowing for presence of co-morbid conditions); repeated clinical, molecular, and imaging data collection performed at a single center; and multi-dimensional analysis to uncover novel biomarkers of PD. A broad biomarker definition will be applied, in line with the recently proposed modular set of biomarker assessments [48]. In addition, participants will be followed for 2 years using a wearable multi-sensor device, which creates the opportunity to continuously monitor aggregated features (e.g., daily and weekly activity level, mean daily pulse rate and its variability, average sleep efficiency per day) to characterize each participant over time. The very liberal inclusion criteria, where participants with co-morbidities are encouraged to join, allow us to collect data from “real-life” patients and explore how co-morbidities impair the overall expression of PD. Moreover, although we aim for a unique dataset, to the extent possible we have harmonized data collection with previously performed cohort studies including the Parkinson Progression Marker Initiative [5], Luxembourg Parkinson’s Study [11] and the Oxford Parkinson Discovery Cohort [12]. In this manner we also support initiatives such as the Critical Path for Parkinson’s

Consortium (CPP), that aim to create integrated unified databases, thus allowing to increase sample sizes or add control populations and further enhance research.

Furthermore, the PPP cohort will contribute data managed through an unprecedented digital security system, which will allow sharing of data with researchers worldwide with maintenance of participants' privacy. The digital security system is based on a multi-point, privacy-by-design strategy: (a) participants provide informed consent, also for the important element of data sharing; (b) signed contractual agreements with researchers are in place to ensure that no attempts towards de-identification or commercialization of the raw data will be attempted; (c) governance policies restrain access to the data to only qualified researchers; (d) an innovative pseudonymization and encryption process is applied [35]; and (e) a protected research environment is used to analyze the data.

Lastly, powered by Verily's analytics capabilities, the PPP will allow us to address research questions of great scientific and clinical value to improve our understanding of PD pathology and variation between patients in terms of disease progression, therapy response (both efficacy and tolerance), and survival. Though we will not be able to adjust for normal age-related changes due to the lack of a matched control group, the PPP data will help to identify new biomarkers to predict differences in prognosis and treatment response between patients, an important step to improving existing treatments, developing new therapeutic approaches, and providing PD patients with a more precise and personalized disease management approach. Finally, this cohort will serve as a source of data for qualified researchers worldwide, allowing them to use their research capacity to further address and enhance the main aims of this study.

2.LITERATURE SURVEY

2.1 Existing problem

Parkinson's disease (PD) is the most common neurodegenerative movement disorder that affects extensive regions of the nervous system. Its current clinical diagnosis is based on motor symptoms that appear late during disease progression when substantial proportions of the nigrostriatal dopaminergic neuron population are lost already. Although disturbances in sleep and other biofunctions often surface years prior to motor impairments and point to a long prodromal phase, these phenotypic signs in a person's midlife lack predictive power. They do, however, signal the unfolding of the disease and suggest molecular correlates that begin deviating early on. Revealing such trajectories, hence, promises not only a better understanding of prodromal PD but may also enable a much-needed earlier diagnosis. A nexus that may harbor such molecular trajectories is the epigenome as key etiological factors of PD—genetics, age, and environment—influence this substrate. An earlier diagnosis would also allow earlier interventions and lifestyle adjustments to improve brain function and reduce symptoms. In this review, we describe the challenges of diagnosing PD early on and highlight the opportunities that may arise from steering research efforts towards comprehensive interrogations of molecular layers during the long-time neglected midlife phase. In particular, we emphasize how existing cohorts of at-risk individuals, available animal models, and suitable markers may come together and aid in revealing molecular trajectories that offer diagnostic utility for PD in its prodromal stage.

2.2 References

- National Institute for Health and Clinical Excellence. Guideline development methods: information for national collaborating centres and guideline developers. London: National Institute for Health and Clinical Excellence; 2004. Ref ID: 20114.

- Parkinson’s Disease Society. Parkinson’s disease: the personal view. London: Parkinson’s Disease Society; 1993. Ref ID: 19909.

- Parkinson’s Disease SocietyOne in twenty: an information pack for younger people with Parkinson’s London: Parkinson’s Disease Society; 2002Ref ID: 19910.

- Parkinson J. Essay on the shaking palsy. London: Sherwood, Neely, and Jones; 1817. Ref ID: 19927.

2.3 Problem Statement Definition

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 present decade of accelerated advances in Medical Sciences, most studies fail to lay focus on ageing diseases. These are diseases that display their symptoms at a much advanced stage and makes a complete recovery almost improbable. Parkinson’s disease (PD) is a member of a larger group of neuromotor diseases marked by the progressive death of dopamine producing cells in the brain. Providing computational tools for Parkinson diseaseusing a set of data that contains medical information is very desirable for alleviating the symptoms that can help the amount of people who want to discover the risk of disease at an early stage. Parkinson’s disease (PD) is the second most commonly diagnosed neurodegenerative disorder of the brain. One could argue, that it is almost incurable and inflicts a lot of pain on the patients. All these make it quite clear that thereis an oncoming need for efficient, dependable and expandable diagnosis of Parkinson’s disease. A dilemma of this intensity requires the automating of the diagnosis to lead accurate and reliable results. It has been observed that most PD Patients demonstrate some sort of impairment in speech or speech dysphonia, which makes speech measurements and indicators one of the most important aspects in prediction of PD. The aim of this work is to develop an effective and accurate model in order to help diagnose the disease accurately at anearlier stage which could in turn help the doctors to assist in the cure and recovery of PD Patients.

Who does the problem affect?	Genetics. Individuals with a parent orsibling who is affected have approximately two times the chance of developing Parkinson's.

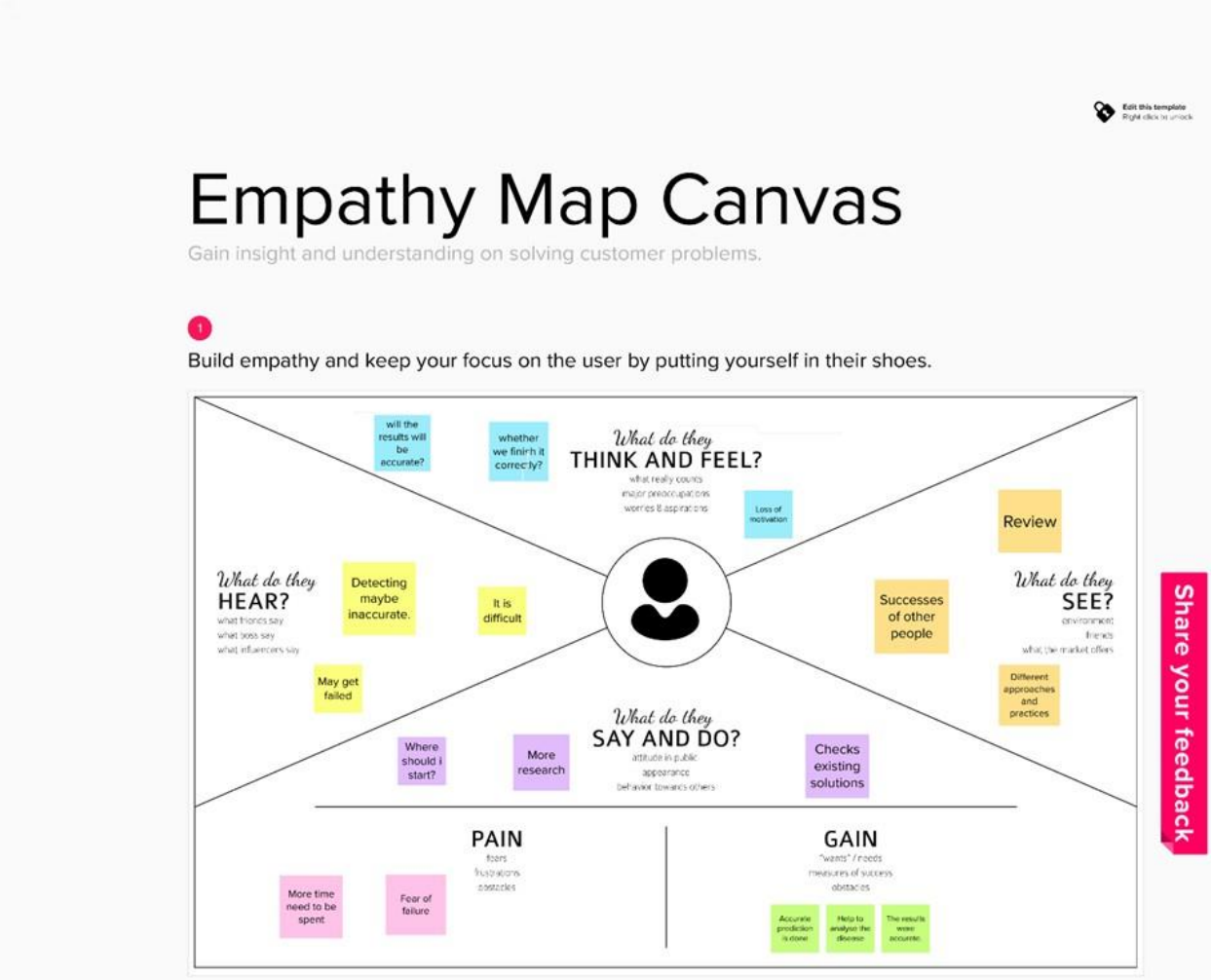
What are the boundaries of the problem?	When nerve cells in the basal ganglia, an area of the brain that controls movement, becomeimpaired and/or die,the person is affected byParkinson’s disease
What is the issue?	Parkinson's disease is a progressive disorder that affects the nervous systemand the parts of the body controlled by the nerves.
When does the issue occur?	Parkinson's disease is caused when lossof nerve cells in the part of the brain called the substantia nigra. Nerve cells in this part of the brainare responsible for producing a chemical called dopamine.
Where doesthe issue occur?	Where doesthe issue occur?
Why is it important that we fix theproblem?	Why is it important that we fix theproblem?

What solution to solve thisissue?	Parkinson's disease can't be cured, but medications can help control the symptoms, often dramatically. In somemore advanced cases, surgery may be advised. Prediction is made to predict the presence of the disease by analysing the factors.
What methodology use to solve theissue?	Machine learning algorithms are usedto predict whether the person has disease or not.

Problem Statement (PS)	**I am (Customer)**	**I'm trying to**	**But**	**Because**	**Which makes mefeel**
PS-1	I am a cutomer.	I am trying to check my disease.	But I unableto know deeply.	Because I am busy in my work.	Which makesme sad.
PS-2	I am a customer.	I am trying to predict my disease.	But I don'tli ke to spend time moreon it.	Because I am busy in my work.	Which makesme little tension.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation& Brainstorming

3.3 Proposed Solution

S.N	Parameter	Description

1	Problem Statement (Problem to besolved)	In the present decadeof accelerated advances in Medical Sciences, most studiesfail to lay focus on ageing diseases. These are diseases that display their symptoms at a much advanced stage and makesa complete recovery almost improbable. Parkinson's disease (PD) is a member of a larger group of neuromotor diseases marked by the progressive death of dopamine producing cells in the brain. Providing computational tools for Parkinson disease using a set of data that contains medicalinformation is very desirable for alleviating the symptoms that can help the amount of people who want to discover the risk of disease at an early stage. Parkinson's disease (PD) is the second most commonly diagnosed neurodegenerative disorder of the brain. One could argue, that it is almost

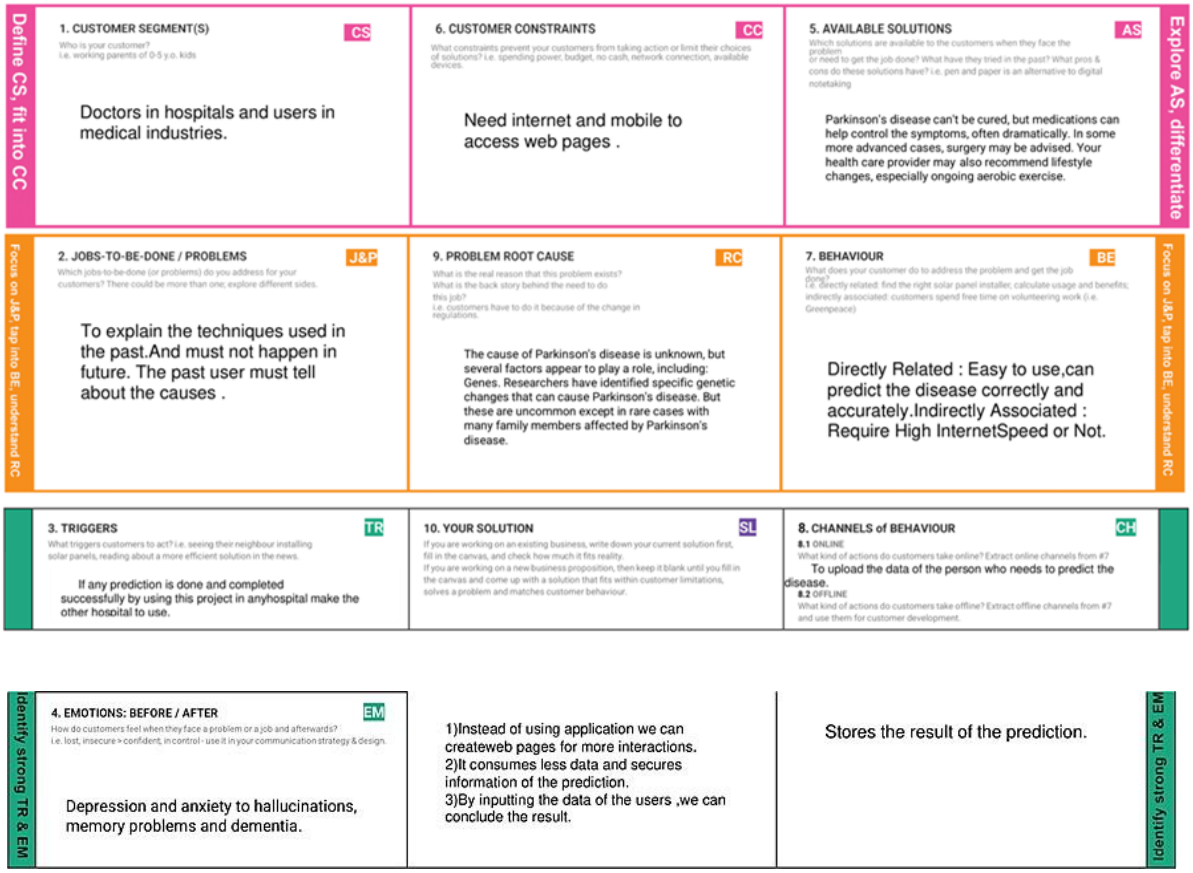
		incurable and inflicts a lot of pain on the patients. All these make it quiteclear that there is an oncoming need for efficient, dependable and expandable diagnosis of Parkinson's disease. A dilemma of this intensity requires the automating of the diagnosis to lead accurate and reliable results. It has been observed that most PD Patients demonstrate some sort of impairment in speech or speech dysphonia, which makes speechmeasurements and indicators one of the most important aspects in prediction of PD. The aim of this work is to develop an effective and accurate model in order to help diagnose the disease accurately at an earlier stage which could in turn help the doctors to assist in the cure and recovery of PD Patients.

2	Idea / Solution description	1. You can directly checkyourprediction. 2. User can apply their data on the respective section to findthe prediction. 3. It willtell whether the personis affected or not with the parkinsons disease.

3	Novelty / Uniqueness	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. Prediction is made topredict the presence of the diseaseby analyzing the factors.
4	Social Impact/ Customer Satisfaction	● Increase salesand conversations ● Personalize the customer experience ● Build brand awareness ● Deal with customer queries ● Accurate and quick prediction.
5	Business Model(Revenue Model)	● Growth opportunities ● Fits into the pocket ● Economical Development ● Uncomplicated interface
6	Scalability of the Solution	● Improved user engagement. ● Drive sales. ● Reduce user acquisition cost. ● Immediate response for user.

3.4 Problem Solution Fit

Project Title: **DETECTING PARKINSONS DISEASE USING MACHINE LEARNING** Project Design Phase-I - Solution Fit Template
Team ID: PNT2022TMID50876



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story/ Sub-Task)
FR-1	User Authentication	The users must be registered first and can be only able to access the web application . This is to ensurethat the web application is used for a good reason.
FR-2	Web Service Management Process	Web Service Management process by WebPortal adminin registering web client to do SSOor member data communication. The web page is hosted in cloud.
FR-3	Data Management	The Web server and Portalmanager can have access to data to edit and updateagain to server.
FR-4	Testing	Applying thealgorithms on the test data
FR-5	Confirmation	Display the result with the description of having Parkinson's or not

4.2 Non – Functional requirements

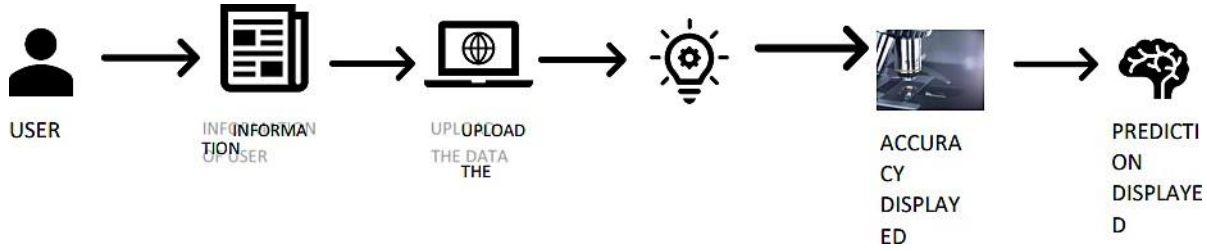
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The webpage loading for users submitting theirinput details at the webapplication must be loaded fast than rendering more time.
NFR-2	**Security**	Authorization access scenarios and definitions, hand-over procedures for patient records. The image andother inputs of patients mustbe highly secured and can't be accessible to others.
NFR-3	**Reliability**	The prediction of the system must be with higheraccuracy so thatthe output fromthe application canbe trusted by the users without any doubts and can be used for further dragonising process with Doctors.
NFR-4	**Performance**	The landing page supporting 5,000 users per hourmust provide 6 second or less response time in a Chromedesktop browser, including the rendering oftext and images and over an LTE connection and theuploading of Data (image) must also should be fast and the outputpage should be rendered within seconds.
NFR - 5	**Availability**	The web application should be available to all Doctors across the globe and can be implemented in every hospital so that the patients can use it effectively
NFR - 6	Scalability	The System must function using Cloud and during a down process also it must satisfy the maximum number of clients. . The system must use higher RAM and CPU processing in Server to handle multiple request at same time

5. PROJECT DESIGN

5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the rightamount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution & Technical Architecture

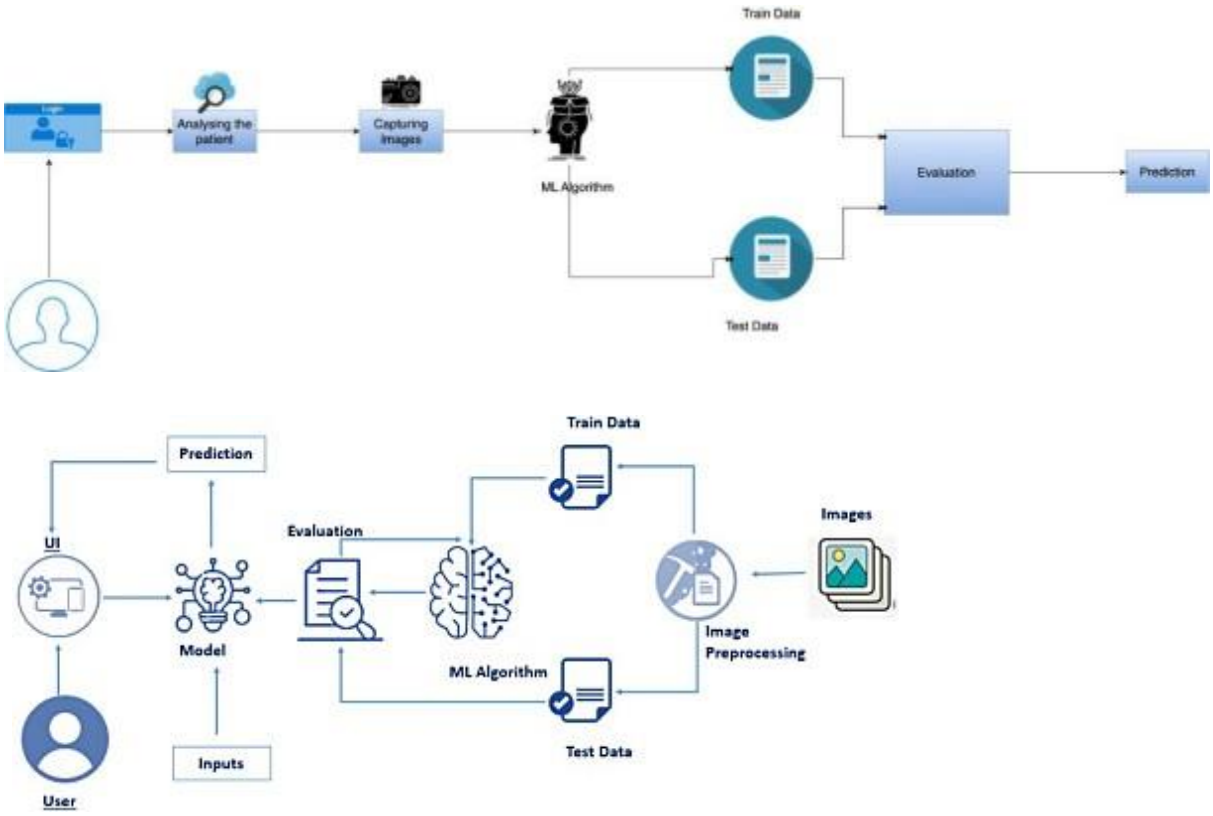


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	user interacts with application e.g. Web UI, Mobile App, etc.	Streamlit python
2.	Application Logic-1	Logic for a process in the application	Python
3.	Application Logic-2	Logic for a process in the application	Google colloboratory
4.	Application Logic-3	Logic for a process in the application	Anaconda,spider
5.	Database	Data Type, Configurations etc.	MySQL
6.	Cloud Database	Database Service on Cloud	Local database
7.	File Storage	File storage requirements	Local Filesystem
8.	External API-1	Purpose of External API used in the application	anaconda, etc.
9.	External API-2	Purpose of External API used in the application	spider, etc.
10.	Machine Learning Model	Purpose of Machine Learning Model	Object Recognition Model, etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Anaconda and spider	Technology of Opensource framework-Anaconda and spider.
2.	Security Implementations	The fire walls used to block the unknown websites.	Firewall

3.	Scalable Architecture	The application works through cloud storage and it can handle maximum number of clients .The system must use higher RAM and CPU processing in server to handle multiple requests at a same time	Technology used
4.	Availability	The application is available in mobile phones,laptop,tablet etc	Technology used
5.	Performance	*The landing page supporting 5,000 users per hour must provide 6 second or less response time in a Chrome desktop browser, including the rendering of text and images and over an LTE connection and the uploading of Data (image) must also should be fast and the output page should be rendered within seconds.*	Technology used

5.3 User Stories

USER TYPE	FUNCTIONAL REQUIREMENT(EPIC)	USER STORY NUMBER	USER STORY /TASK	ACCEPTANCE CRITERIA	PRIORITY	RELEASE
CUSTOMER(PATIENT)	UPLOADS DATA	USN-1	USER UPLOADS DATA THROUGH WEBSITES	TYPE OF DATA IS ACCEPTED	HIGH	SPRINT-1
CUSTOMER(DOCTOR)	UPLOADS DATA	USN-1	USER UPLOADS DATA THROUGH WEBSITES	TYPE OF DATA IS ACCEPTED	HIGH	SPRINT-1

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement(Epic)	User Story Number	User Story/ Task	Story Points	Priority	Team Members
Sprint -1	REGISTRATION	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	high	3
Sprint -1		USN-2	As a user, I will receivea confirmation emailonce I haveregistered for theapplication	1	high	1
Sprint -2		USN-3	As a user, I can register for the application through Facebook	2	low	1
Sprint -1		USN-4	As a user, I can register for the application through google account	2	high	1
Sprint -1	LOGIN	USN-5	As a user, I can log into theapplication by entering email & password	1	high	1

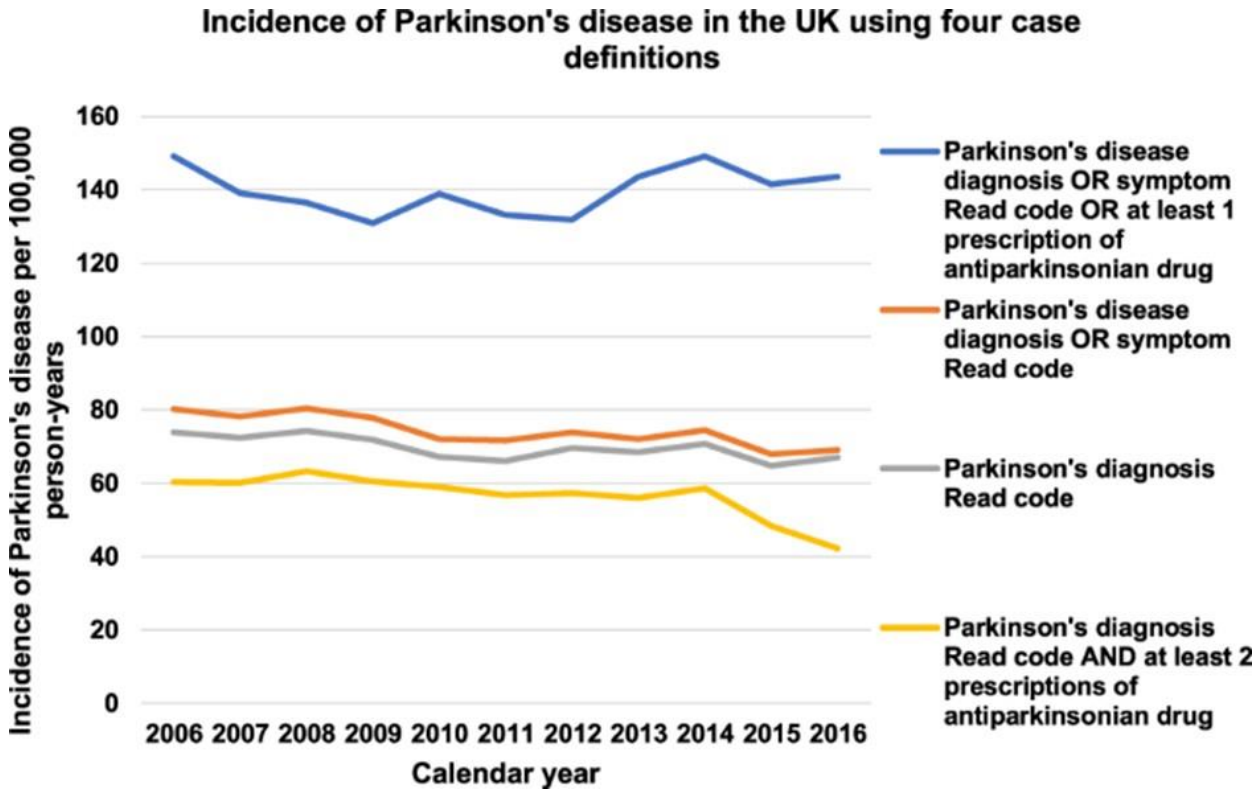
Sprint-1	DASHBOARD	USN-6	As a customer I can checkwith all theclothes available on the website andchoose the ones which I want	3	high	3
Sprint-4	Customer support	USN-7	As a user I want to contact with the customer supportwhen there is any querywith the the application	2	low	2
Sprint-1	User details display	USN-8	As a customer I should be able to seeall my given details filled with the registration process	2	high	2
Sprint-2	algorithm	USN-9	As a customer, I should be able to getthe perfect prescription for the diseasecriticality and get the correct doctor details	5	high	4
Sprint-2		USN-10	As a customer I should be updated withvarious best available	2	medium	2
Sprint-3	IBM watson for storage and organization	USN-11	As a customer I should be able to givemy images and predict the out come and the prescription for my disease criticality	3	high	2
Sprint-4	Cart management	USN-12	As a customer I can add and manage all the chosen products and place myorder	5	high	4

6.2 Sprint Delivery Schedule

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project & gatheringinformation by referring the,technical papers,research publications etc.	28 SEPTEMBER 2022
Prepare Empathy Map	Prepare Empathy Map Canvasto capture the user Pains & Gains, Prepare list of problemstatements	24 SEPTEMBER 2022
Ideation	List the by organizing the brainstorming session andprioritize the top 3 ideas based on the feasibility & importance.	25 SEPTEMBER 2022
Proposed Solution	Prepare the proposed solutiondocument, which includes thenovelty, feasibility of idea, business model, social impact, scalability of solution, etc.	23 SEPTEMBER 2022
Problem Solution Fit	Prepare problem - solution fitdocument.	30 SEPTEMBER 2022
Solution Architecture	Prepare solution architecturedocument.	28 SEPTEMBER 2022

Customer Journey	Prepare the customer journeymaps to understand the user interactions & experiences with the application (entry to exit).	20 OCTOBER 2022
Functional Requirement	Prepare the functional requirement document.	8 OCTOBER 2022
Data Flow Diagrams	Draw the data flow diagrams and submitfor review.	9 OCTOBER 2022
Technology Architecture	Prepare the technology architecture diagram.	10 OCTOBER 2022
Prepare Milestone & ActivityList	Prepare the milestones& activity list of the project.	22 OCTOBER 2022
Project Development - Delivery of Sprint-1, 2, 3 & 4	Develop & submit the developed code by testing it.	17 NOVEMBER 2022

6.3 Report from JIRA



7. CODING & SOLUTIONING

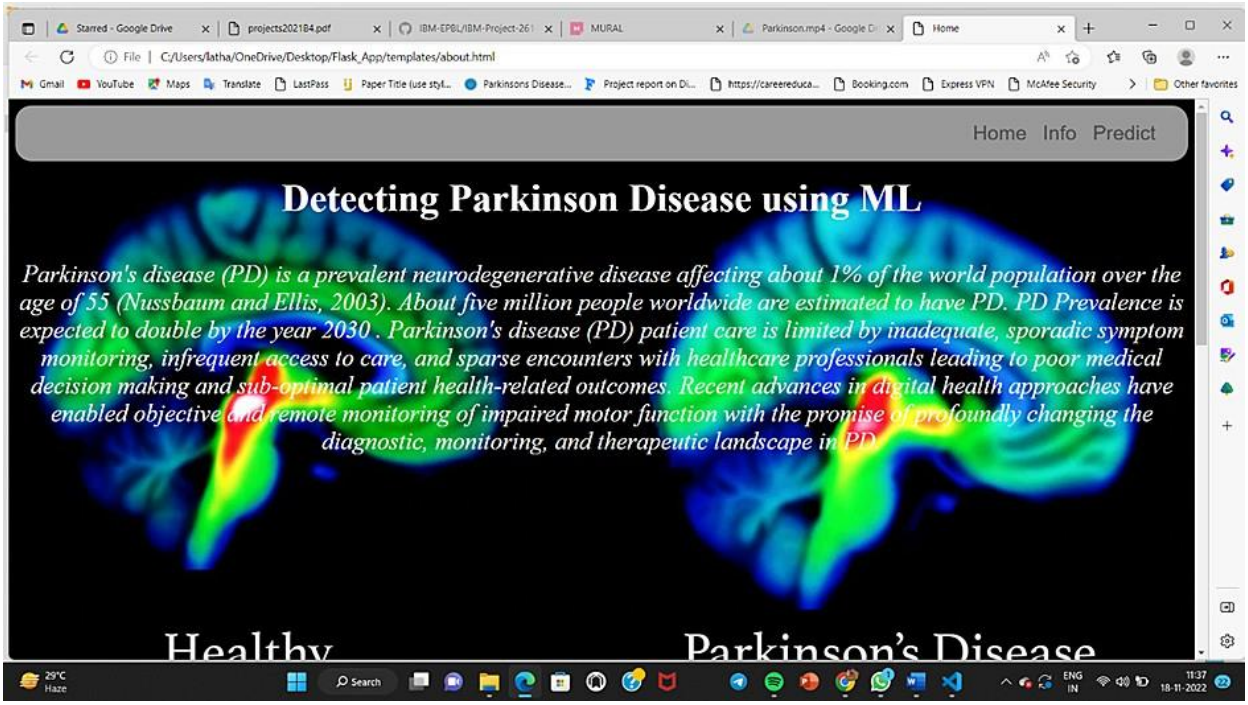
7.1 Feature 1

The home page provides details about the parkinsons disease.

```
<!DOCTYPE html>
<html>
<head>
<title>Home</title>
<style>
body
{
   background-image:
url("https://www.telegraph.co.uk/content/dam/science/2019/06/19/TELEMMGLPICT000201206329_trans_N
```

```
vBQzQNjv4BqHYnrW5_va2i5YQ7jbSeHyuC4yxA4SBURrqE8SKy5rSc.jpeg");
    background-size: cover;
}
.pd{
padding-bottom:100%;}
.navbar
{
margin: 0px;
padding:20px;
background-color:white;
opacity:0.6;
color:black;
font-family:'Roboto',sans-serif;
font-style: italic;
border-radius:20px;
font-size:25px;
}
a
{
color:grey;
float:right;
text-decoration:none;
font-style:normal;
padding-right:20px;
}
a:hover{
background-color:black;
color:white;
border-radius:15px;
font-size:30px;
padding-left:10px;
}
p
{
color:white;
font-style:italic;
font-size:30px;
}
</style>
```

```
</head>
<body>
<div class="navbar">
<a href="/upload" >Predict</a>
<a href="/info">Info</a>
<a href="/about">Home</a>
<br>
</div>
<br>
<center><b class="pd"><font color="white" size="15" font-family="Comic Sans MS" >Detecting Parkinson Disease using ML</font></b></center>
<div>
<br>
<center>
<p>Parkinson's disease (PD) is a prevalent neurodegenerative disease affecting about 1% of the world population over the age of 55 (Nussbaum and Ellis, 2003). About five million people worldwide are estimated to have PD. PD Prevalence is expected to double by the year 2030 . Parkinson's disease (PD) patient care is limited by inadequate, sporadic symptom monitoring, infrequent access to care, and sparse encounters with healthcare professionals leading to poor medical decision making and sub-optimal patient health-related outcomes. Recent advances in digital health approaches have enabled objective and remote monitoring of impaired motor function with the promise of profoundly changing the diagnostic, monitoring, and therapeutic landscape in PD.</p>
</center>
</div>
</body>
</html>
```



7.2 Feature 2

The info page provides the information about the parkinsons disease.

```
<!DOCTYPE html>
<html>
<head>
<style>
.navbar
{
margin: 0px;
padding:20px;
background-color:white;
opacity:0.6;
color:black;
font-family:'Roboto',sans-serif;
font-style: italic;
border-radius:20px;
font-size:25px;
}
a
{
```

```
color:grey;
float:right;
text-decoration:none;
font-style:normal;
padding-right:20px;
}
a:hover{
background-color:black;
color:white;
border-radius:15px;
font-size:30px;
padding-left:10px;
}

img{
width:50px;
height:700px;
padding:10px;
margin-top:0px;
}
img:hover{
border-radius:100px;
border-color:grey;
border-shadow:10px;

}
body{
background-color:black;

}
h1{
font-size:60px;
text-align:center;
color:white;
font-style:italic;
font-weight:bolder;
}
div{
margin-left:50px;
```

```
}
img{
width:500px;
height:400px;
padding:10px;
margin-top:0px;
}
img:hover{
border-radius:100px;
border-color:grey;
border-shadow:10px;

}
</style>
<title>Info</title>
</head>
<body>
<div class="navbar">
<a href="/upload" >Predict</a>
<a href="/info">Info</a>
<a href="/about">Home</a>
<br>
</div>
<div>
<h1>Prevention is better than cure!</h1>
<div>

<span><img src="https://i.pinimg.com/originals/98/c0/97/98c097bab991f995236c2331615fae1a.jpg"
title="Disease"></span>
<span><img src="https://www.connectneurophysiotherapy.com/wp-content/uploads/2019/05/hoehn-and-
Yahr.png" title="Stages"></span>

<div>
</body>
</html>
```

8. TESTING

8.1 Test Cases

As is evident from the code itself,

- load_data_from_path() function will load data from a given file path and return the data frame
- load_insurance_data() function will load the data we have downloaded from Kaggle and return a data frame
- load_processed_insurance_data() function will load the processed dataset from the hardcoded path

Function for splitting training and testing data:

```
def split_train_test_data():
    df = load_processed_insurance_data()
    X = df.drop(columns=["charges"])
    # In the above line, the column sex is also dropped but let's see what's the effect of keeping sex on the
    # invariance test
    y = df["charges"]
```

```
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
    save_train_test_data(X_train, y_train, X_test, y_test)
    return X_train, y_train, X_test, y_test
```

The above function will load the processed insurance data first. Then it will remove the column to be predicted from it (X) and create a separate data frame with output charge values (y). Next, it will split the X and y data frames so formed into training and testing datasets and return those to the caller.

Functions for saving the dataset:

```
def save_insurance_data(df: pd.DataFrame):
    output_path = "data/processed/processed_insurance_data.csv"
    df.to_csv(output_path, index=False)
    return
```

The above function will save the processed data to the folder named “processed”

under the data folder for further use.

```
def save_train_test_data(Xtrain: pd.DataFrame, Ytrain: pd.DataFrame, Xtest: pd.DataFrame, Ytest: pd.DataFrame):
    Xtrain.to_csv("data/model_training_data/training_data.csv", index=False)
    Ytrain.to_csv("data/model_training_data/training_data_result.csv", index=False)
    Xtest.to_csv("data/model_testing_data/testing_data.csv", index=False)
    Ytest.to_csv("data/model_testing_data/testing_data_result.csv", index=False)
    return
```

The above function will save the previously split training and testing data to their designated locations.

Next, let’s write the script for pre-processing the data.

Data Pre-Processing Script

In data pre-processing, we will load the insurance data using the function load_insurance_data() as discussed before. We will introduce two new columns in the dataset:

- age_range: this column will categorize each row into an age range. The categorization is as follows: all data points with
    - age value less than 30 will have value for age_range as 1;
    - age value less than 40 but greater than or equal to 30 will have value for age_range as 2
    - age value less than 50 but greater than or equal to 40 will have value for age_range as 3
    -
    - age value greater than 50 will have value for age_range as 4
- have_children: this column will have value as “Yes” for all rows with values for children feature as greater than 0 and “No” value for the rest

```
def data_preprocessing():
    """ Runs data processing scripts to turn raw data from (../raw) into
        cleaned data ready to be analyzed (saved in ../processed).
    """
    logger = logging.getLogger(_name_)
    logger.info('making final data set from raw data')
    # df = pd.DataFrame()
    df = load_insurance_data()
    # creating new feature by using age column
    df["age_range"] = 1000
    for i in range(len(df["age"])):
```

```
    if df["age"][i] < 30:
        df["age_range"][i] = 1
    elif 30 <= df["age"][i] < 40:
        df["age_range"][i] = 2
    elif 40 <= df["age"][i] < 50:
        df["age_range"][i] = 3
    elif df["age_range"][i] >= 50:
        df["age_range"][i] = 4
df["have_children"] = ["No" if i == 0 else "Yes" for i in df["children"]]
```

```
cat_variable = ['sex', 'smoker', 'region', 'have_children'] lb = LabelEncoder()

df[cat_variable] = df[cat_variable].apply(lambda col:

lb.fit_transform(col.astype(str)))


save_insurance_data(df)
return
```

After introducing two new columns in the dataset, we will encode the columns with categorical values using sklearn's LabelEncoder. A label encoder transforms the labels of a categorical column into numeric form for efficient processing by the machine learning algorithms.

After the column transformation, we save the processed dataset using the save_insurance_data() discussed before.

Next, we will write train_model.py and predict_model.py scripts.

Training and Prediction Scripts

For training, we will consider only two models: Linear Regression and K Nearest Neighbors.

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.neighbors import KNeighborsRegressor

def linear_regression(xtrain, ytrain):

    lr = LinearRegression()

    lr.fit(xtrain, ytrain)

    return lr

def k_neighbours(xtrain, ytrain):

    knn = KNeighborsRegressor()

    knn.fit(xtrain, ytrain)

    return knn
```

As is known from the code above, linear_regression function will initiate the LinearRegression() from the sklearn package and train it on the given training data. k_neighbours() function behaves similarly.

Let's look at the prediction script now:

```
import pandas as pd

def predict_on_test_data(model, xtest):

    y_test = model.predict(xtest)

    filename = str(model.__class__.__name__)+"predicted output.csv"

    prediction = pd.DataFrame(y_test)

    pd.DataFrame(y_test).to_csv("data/model_testing_data/"+filename)

    return prediction
```

predict_on_test_data() function will simply take the model passed into it and predict the values for the test data passed as a second argument. Then it will save the resulting predictions into the appropriate folders and return the prediction data frame.

Let's write the pre-train and post-train test scripts now.

Pre-train and Post-train Test Scripts

Before writing the tests, let us write fixtures. We will write fixtures using the

decorator @pytest.fixtures. Fixtures are designed to run before each of the test functions that it is applied to. They provide important context to the tests. Refer to this link for more information on pytest fixtures

```
@pytest.fixture
def data_preparation():
    data_preprocessing()
    return split_train_test_data()
@pytest.fixture
def linear_regression_prediction(data_preparation):
    xtrain, ytrain, xtest, ytest = data_preparation
    lr = linear_regression(xtrain, ytrain)
    ypred = predict_on_test_data(lr, xtest)
    return xtest, ypred
@pytest.fixture
def k_neighbors_prediction(data_preparation):
    xtrain, ytrain, xtest, ytest = data_preparation
    knn = k_neighbours(xtrain, ytrain)
    ypred = predict_on_test_data(knn, xtest)
    return xtest, ypred
```

```
@pytest.fixture
def return_models(data_preparation):
    xtrain, ytrain, xtest, ytest = data_preparation
    lr = linear_regression(xtrain, ytrain)
    knn = k_neighbours(xtrain, ytrain)
```

```
    return [lr, knn]
```

We've designed 4 fixtures:

- data_preparation(): this fixture will call the data_preprocessing() function defined before to pre-process the raw data. Next, it will call the split_train_test_data() function to split the data and pass it to the caller
- linear_regression_prediction(): this fixture will take data_preparation() fixture values as it's input. It will train the linear_regression model on the training data passed by the data_preparation fixture followed by calling predict_on_test_data() to predict the values based on the trained model and will finally return the test data and predicted values
- k_neighbors_prediction(): this fixture will work in the same way as the previous fixture except that it will train the model using the K Nearest Neighbors algorithm
- return_models(): this fixture will take data-preparation as its input and return the trained linear regression and k nearest neighbor models on the training data.

Machine Learning Pre-train Test Script

We will check for two things in this example – test data leakage and predicted output shape validation. The script is as follows:

```
import pytest_check as check


def test_data_leak(data_preparation):
    xtrain, ytrain, xtest, ytest = data_preparation
    concat_df = pd.concat([xtrain, xtest])
    concat_df.drop_duplicates(inplace=True)
```

```
    assert concat_df.shape[0] == xtrain.shape[0] + xtest.shape[0]
def test_predicted_output_shape(linear_regression_prediction, k_neighbors_prediction):
    print("Linear regression")
    xtest, ypred = linear_regression_prediction
    check.equal(ypred.shape, (xtest.shape[0], ))
    # assert ypred.shape == (xtest.shape[0], 1)
    print("K nearest neighbours")
    xtest, ypred = k_neighbors_prediction
    check.equal(ypred.shape, (xtest.shape[0], ))
    # assert ypred.shape == (xtest.shape[0], )
```

In the test_data_leak() function, we take the data_preparation() fixture as the input which will provide us with the dataset. Next, we will concatenate the training and test data and drop the duplicates from it. Now, we'll check whether the size of the concatenated datasets without duplicates is the same as the addition of shapes of training and testing data.

In the test_predicted_output_shape() function, we take the linear_regression_prediction() and k_neighbors_prediction() fixtures discussed before as the input. Next, we obtain their prediction values and compare their shape with that of the testing dataset. Here, for illustration purposes, the wrong comparison is done. The actual shape comparison should be (xtest.shape[0], 1).

Also note that, instead of assert, we have pytest_check. It acts as a soft assert meaning that even if an assertion fails in between, the test will not stop there but instead proceed till completion and report all the assertion fails.

8.2 User Acceptance Testing

Number	Acceptance Requirement	Critical		Test Result		Comments
		Yes	No	Accept	Reject	
1	The system must execute to end of job.	√				Payroll will not run in a production status until this requirement has been met.
2	The results of payroll must be correct.	√				Payroll will not run in a production status until this requirement has been met.

Test Execution

Usually, when possible, this testing happens in a conference or a war room sort of a set up where the users, PM, QA team representatives all sit together for a day or two and work through all the acceptance test cases.

Or in case of the QA team performing the tests, we run the test cases on the AUT.

Once all the tests are run and the results are in hand, the **Acceptance Decision** is made. This is also called the **Go/No-Go decision**. If the users are satisfied it's a Go, or else it's a No-go.
Reaching the acceptance decision is typically the end of this phase.

Tools & Methodologies

Typically, the type of software tools that are used during this testing phase is similar to the tools used while performing functional testing.

9. RESULTS

9.1 Performance Metrics

Evaluation is a process during the development of the model to check whether the model is the best fit for the given problem and corresponding data.

Classification Evaluation Metrics:

These model evaluation techniques are used to find out the accuracy of models built in the classification type of machine learning models. We have three types of

evaluation methods.

- Accuracy_score
- Confusion matrix
- Roc- Auc Curve

Confusion Matrix

It is a matrix representation of the results of any binary testing.

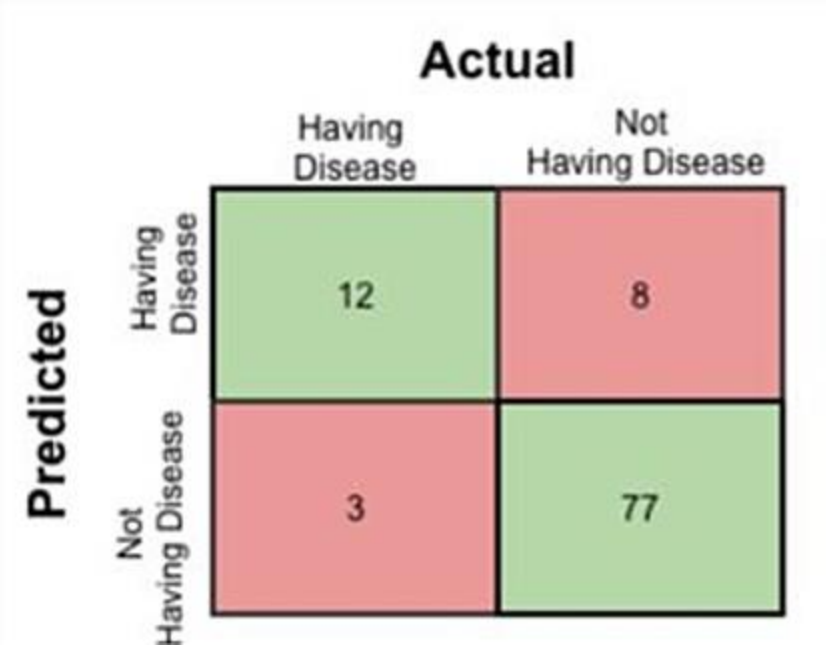


Fig: Confusion Matrix of prediction of a disease

1. True Positive: 12 (You have predicted the positive case correctly!)
2. True Negative: 77 (You have predicted negative case correctly!)
3. False Positive: 8 (You have predicted these people as having disease, but in actual they do not have.)
4. False Negative: 3 (Wrong predictions)

We can use the **predict** method on the model and pass **X_test** as a parameter to get

the output as predictions.

The output of the confusion matrix and accuracy are as follows.

10. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

1. No infection risk associated with surgical thalamotomy or surgical implants

2. "Test" sonications to confirm precise targeting prior to full ablation

3. When successful, the results are immediately evident.

DISADVANTAGES:

A personalized medicine approach that treats an individual with PD in a timely manner with the optimal treatment requires understanding the enormously complex and diverse set of factors that contribute to PD. The disease processes that lead to PD involve numerous potential variables and pathways operating at cellular and molecular levels. Most of these processes unfold over the course of many years and begin well before individuals start having symptoms. People with PD may also differ significantly in terms of the symptoms they experience, the severity of those symptoms, disease progression, and their response to treatment and risk of complications.

Improving our understanding of what causes the complexity and diversity of PD is a major challenge for researchers. Tools are needed to group people with similar types of PD so that individuals who are most likely to benefit from clinical trials can be studied and their responses to treatment can be compared in a meaningful way.

11. CONCLUSION

Parkinson's is the second most neurodegenerative disease which has no cure.

It results in difficulty of body movements, anxiety, breathing problems, loss of smell, depression, and speech. In this paper, the three different machine learning algorithms used to measure the performance are KNN, Naïve Bayes, and Logistic Regression applied on the dataset. The author chose the voice features of patients as the dataset contains more than 700 features and finally took the ten important features that are useful to evaluate the system. The author compared all the three machine learning methods accuracies and based on this one prediction model is generated. Hence, the aim is to use various evaluation metrics like confusion matrix, accuracy, precision, recall, and f1-score which predicts the disease efficiently. Comparing all the three the Naïve Bayes classifier gives the highest accuracy of 81%.The diagnosis of Parkinson's Disease is not direct which means that one particular test like blood test or ECG cannot determine whether a person is suffering from PD or not. Doctors go through the medical history of a patient, followed by a thorough neurological examination. They find out at least two cardinal symptoms among the subjects and then predict whether the subject is suffering from PD. The misdiagnosis rate of PD is significant due to a no definitive test. In such case it will be helpful for us to aid the doctor by providing a machine learning model. The prediction models are developed using machine learning techniques of boosted logistic regression, classification trees , Bayes Net and multilayer perceptron based on these significant features. It is observed that the performance is better. It is demonstrated that Boosted Logistic Regression produce superior results. These results encourage us to try other ensemble learning techniques. The present work employs different machine learning algorithms which are not used in [3]. This study plays an important role in having a comparative analysis of various machine learning algorithms. In conclusion, this models can provide the nuclear experts an assistance that can aid them in better and accurate decision making and clinical diagnosis. It is also found that the proposed method is fully automated and provides improved performance and hence can be recommended for real life applications.

12. FUTURE SCOPE

In future, these models can be trained with different datasets that have best features and can be predicted more accurately. If the accuracy rate increases, it can be used by the laboratories and hospitals so that it is easy to predict in early stages.

This models can be also used with different medical and disease datasets. In future the work can be extended by building a hybrid model that can find more than one disease with an accurate dataset and that dataset has common features of two diseases. In future the work can extended to build a model that may extract more important features among all features in the dataset so that it produce more accuracy.

13. APPENDIX

SOURCE CODE

main.css

```
.img-preview {
    width: 256px;
    height: 256px;
    position: relative;
    border: 5px solid #F8F8F8;
    box-shadow: 0px 2px 4px 0px rgba(0, 0, 0, 0.1);
    margin-top: 1em;
    margin-bottom: 1em;
}

.img-preview>div {
     width: 100%;
    height: 100%;
    background-size: 256px 256px;
    background-repeat: no-repeat;
    background-position: center;
}

input[type="file"] {
    display: none;
}

.upload-label{
    display: inline-block;
    padding: 12px 30px;
    background: #39D2B4;
    color: #fff;
```

```
    font-size: 1em;
    transition: all .4s;
    cursor: pointer;
}

.upload-label:hover{
    background: #34495E;
    color: #39D2B4;
}

.loader {
    border: 8px solid #f3f3f3; /* Light grey */
    border-top: 8px solid #3498db; /* Blue */
    border-radius: 50%;
    width: 50px;
    height: 50px;
    animation: spin 1s linear infinite;
}

@keyframes spin {
    0% { transform: rotate(0deg); }
    100% { transform: rotate(360deg); }
}
```

main.js

```
$(document).ready(function () {
    // Init
    $('.image-section').hide();
    $('.loader').hide();
    $('#result').hide();

    //        Upload        Preview
    function readURL(input) {
        if (input.files && input.files[0]) {
            var reader = new FileReader();
            reader.onload = function (e) {
```

```
            $('#imagePreview').css('background-image', 'url(' + e.target.result + ')');
            $('#imagePreview').hide();
            $('#imagePreview').fadeIn(650);
        }
        reader.readAsDataURL(input.files[0]);
    }
}
$("#imageUpload").change(function () {
    $('.image-section').show();
    $('#btn-predict').show();
    $('#result').text('');
    $('#result').hide();
    readURL(this);
});

// Predict
$('#btn-predict').click(function () {
    var form_data = new FormData($('#upload-file')[0]);

    // Show loading animation
    $(this).hide();
    $('.loader').show();

    // Make prediction by calling api /predict
    $.ajax({
        type: 'POST',
        url: '/predict',
        data: form_data,
        contentType: false,
        cache: false,
        processData: false,
        async: true,
        success: function (data) {
            // Get and display the result
            $('.loader').hide();
            $('#result').fadeIn(600);
            $('#result').text('Prediction : '+data);
            console.log('Success!');
        },
```

```
        });
    });

});
```

about.html

```
<!DOCTYPE html>
<html>
<head>
<title>Home</title>
<style>
body
{
    background-image:
url("https://www.telegraph.co.uk/content/dam/science/2019/06/19/TELEMMGLPICT000201206329_trans_NvBQzQNjv4BqHYnrW5_va2i5YQ7jbSeHyuC4yxA4SBURrqE8SKy5rSc.jpeg");
    background-size: cover;
}
.pd{
padding-bottom:100%;}
.navbar
{
margin: 0px;
padding:20px;
background-color:white;
opacity:0.6;
color:black;
font-family:'Roboto',sans-serif;
font-style: italic;
border-radius:20px;
font-size:25px;
}
a
{
color:grey;
float:right;
text-decoration:none;
```

```
font-style:normal;
padding-right:20px;
}
a:hover{
background-color:black;
color:white;
border-radius:15px;
font-size:30px;
padding-left:10px;
}
p
{
color:white;
font-style:italic;
font-size:30px;
}
</style>
</head>
<body>
<div class="navbar">
<a href="/upload" >Predict</a>
<a href="/info">Info</a>
<a href="/about">Home</a>
<br>
</div>
<br>
<center><b class="pd"><font color="white" size="15" font-family="Comic Sans MS" >Detecting Parkinson Disease using ML</font></b></center>
<div>
<br>
<center>
<p>Parkinson's disease (PD) is a prevalent neurodegenerative disease affecting about 1% of the world population over the age of 55 (Nussbaum and Ellis, 2003). About five million people worldwide are estimated to have PD. PD Prevalence is expected to double by the year 2030 . Parkinson's disease (PD) patient care is limited by inadequate, sporadic symptom monitoring, infrequent access to care, and sparse encounters with healthcare professionals leading to poor medical decision making and sub-optimal patient health-related outcomes. Recent advances in digital health approaches have enabled objective and remote monitoring of impaired motor function with the promise of profoundly changing the diagnostic, monitoring, and therapeutic landscape in PD.</p>
```

```
</center>
</div>
</body>
</html>
```

base.html

```
<html lang="en">

<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <meta http-equiv="X-UA-Compatible" content="ie=edge">
    <title>Predict</title>
    <link href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css" rel="stylesheet">
    <script src="https://cdn.bootcss.com/popper.js/1.12.9/umd/popper.min.js"></script>
    <script src="https://cdn.bootcss.com/jquery/3.3.1/jquery.min.js"></script>
    <script src="https://cdn.bootcss.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>
    <link href="{{ url_for('static', filename='css/main.css') }}" rel="stylesheet">
<style>
.bar
{
margin: 0px;
padding:20px;
background-color:white;
opacity:0.6;
color:black;
font-family:'Roboto',sans-serif;
font-style: italic;
border-radius:20px;
font-size:25px;
}
a
{
color:grey;
float:right;
text-decoration:none;
```

```
font-style:normal;
padding-right:20px;
}
a:hover{
background-color:black;
color:white;
border-radius:15px;
font-size:30px;
padding-left:10px;
}
body
{
   background-image:
url("https://www.elderresearch.com/hubfs/BLOG_Parkinson%E2%80%99s%20Test%20Recommendation%20
Engine.jpg");
   background-size: cover;

}
</style>
</head>

<body>

<div class="bar">
<a href="/upload" >Predict</a>
<a href="/info">Info</a>
<a href="/about">Home</a>
<br>
</div>
   <nav class="navbar navbar-dark bg-dark">
      <div class="container">
         <a class="navbar-brand" href="#">Parkinson Detection</a>
      </div>
   </nav>
   <div class="container">
      <center> <div id="content" style="margin-top:2em">{% block content %}{% endblock
%}</div></center>
   </div>
</body>
```

```
<footer>
   <script src="{{ url_for('static', filename='js/main.js') }}" type="text/javascript"></script>
</footer>

</html>
```

index6.html

```
{% extends "base.html" %} {% block content %}

<h2><center>Parkinson Classifier</center></h2>

<div>
   <form id="upload-file" method="post" enctype="multipart/form-data">
   <center>    <label for="imageUpload" class="upload-label">
          Choose...
       </label>
       <input type="file" name="file" id="imageUpload" accept=".png, .jpg, .jpeg">
   </center></form>

  <center> <div class="image-section" style="display:none;">
       <div class="img-preview">
          <div id="imagePreview">
          </div></center>
       </div>
       <center><div>
          <button type="button" class="btn btn-primary btn-lg " id="btn-predict">Predict!</button>
       </div></center>
   </div>

   <div class="loader" style="display:none;"></div>

   <h3 id="result">
       <span> </span>
   </h3>
```

```
</div>

</div>



{% endblock %}
```

info.html

```
<!DOCTYPE html>
<html>
<head>
<style>
.navbar
{
margin: 0px;
padding:20px;
background-color:white;
opacity:0.6;
color:black;
font-family:'Roboto',sans-serif;
font-style: italic;
border-radius:20px;
font-size:25px;
}
a
{
color:grey;
float:right;
text-decoration:none;
font-style:normal;
padding-right:20px;
}
a:hover{
background-color:black;
color:white;
border-radius:15px;
```

```
font-size:30px;
padding-left:10px;
}

img{
width:50px;
height:700px;
padding:10px;
margin-top:0px;
}
img:hover{
border-radius:100px;
border-color:grey;
border-shadow:10px;

}
body{
background-color:black;

}
h1{
font-size:60px;
text-align:center;
color:white;
font-style:italic;
font-weight:bolder;
}
div{
margin-left:50px;
}
img{
width:500px;
height:400px;
padding:10px;
margin-top:0px;
}
img:hover{
border-radius:100px;
border-color:grey;
```

```
border-shadow:10px;

}
</style>
<title>Info</title>
</head>
<body>
<div class="navbar">
<a href="/upload" >Predict</a>
<a href="/info">Info</a>
<a href="/about">Home</a>
<br>
</div>
<div>
<h1>Prevention is better than cure!</h1>
<div>

<span><img src="https://i.pinimg.com/originals/98/c0/97/98c097bab991f995236c2331615fae1a.jpg"
title="Disease"></span>
<span><img src="https://www.connectneurophysiotherapy.com/wp-content/uploads/2019/05/hoehn-and-
Yahr.png" title="Stages"></span>

<div>
</body>
</html>
```

app.py

```
import pickle
import cv2
from skimage import feature
from flask import Flask,request, render_template
import os.path
app=Flask(__name__)#our flask app

@app.route("/") #default route
def about():
    return render_template("about.html")#rendering html page
```

```
@app.route("/about") #route about page
def home():
    return render_template("about.html")#rendering html page

@app.route("/info") # route for info page
def information():
    return render_template("info.html")#rendering html page

@app.route("/upload") # route for uploads
def test():
    return render_template("index6.html")#rendering html page

@app.route('/predict', methods=['GET', 'POST'])
def upload():
    if request.method == 'POST':
        f=request.files['file'] #requesting the file
        basepath=os.path.dirname(_file_)#storing the file directory
        filepath=os.path.join(basepath,"uploads",f.filename)#storing the file in uploads folder
        f.save(filepath)#saving the file
        #Loading the saved model
        print("[INFO] loading model...")
        model = pickle.loads(open('parkinson.pkl', "rb").read())
        # pre-process the image in the same manner we did earlier
        image = cv2.imread(filepath)
        output = image.copy()
        # load the input image, convert it to grayscale, and resize
        output = cv2.resize(output, (128, 128))
        image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
        image = cv2.resize(image, (200, 200))
        image = cv2.threshold(image, 0, 255,
        cv2.THRESH_BINARY_INV | cv2.THRESH_OTSU)[1]


        # quantify the image and make predictions based on the extracted
        # features using the last trained Random Forest
        features = feature.hog(image, orientations=9,
        pixels_per_cell=(10, 10), cells_per_block=(2, 2),
        transform_sqrt=True, block_norm="L1")
```

```
        preds = model.predict([features])
        print(preds)
        ls=["healthy","parkinson"]
        result = ls[preds[0]]



        # draw the colored class label on the output image and add it to
        # the set of output images
        color = (0, 255, 0) if result == "healthy" else (0, 0, 255)
        cv2.putText(output, result, (3, 20), cv2.FONT_HERSHEY_SIMPLEX, 0.5,color, 2)
        cv2.imshow("Output", output)
        cv2.waitKey(0)
        return result
    return None

if __name__=="__main__":
    #app.run(debug=False)#running our app
    app.run(host='0.0.0.0', port=8000,debug=False)
```

GITHUB & PROJECT DEMO LINK

PROJECT DEMO LINK: https://drive.google.com/file/d/1bZfV-dIFbGNFRHq5ozWfTZIyxL68oFQl/view?usp=drivesdk

GITHUB LINK: https://github.com/IBM-EPBL/IBM-Project-39528-1660454891