JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING BASED SYSTEM FOR PREDICTING PARKINSON'S DISEASE

[1]V. Neelima Devi, [2] Rampi Sireesha,
[1]Head of Department, [2] MCA 2nd year,
[1]Master of Computer Application,
[1]Sanketika Vidya Parishad Engineering College, Visakhapatnam, India

ABSTRACT

Parkinson's disease is a disease which is a disorder in the nervous system. This disease affects the movement of the human body. In today's world, around 1 million people are suffering from this disease. This is a disorder which produces neurodegenerative dopamine-producing neurons in the brain. The project is to help to detect Parkinson's disease by symptoms in the human body. The project will be made by a new machine learning algorithm.

It is one of the type of diseases that display their symptoms at a much- advanced stage and makes a complete recovery almost improbable. Parkinson's disease (PD) is the second most commonly diagnosed neurodegenerative disorder of the brain. It is almost incurable and inflicts a lot of pain on the patients. All these make it quite clear that there is an oncoming need for efficient, dependable and expandable diagnosis of Parkinson's disease.

KEYWORDS: Biomarkers;Ferroptosis;Genetics;Gut-brainaxis;Neuroinflammation; **Parkinson's disease**; Sleep disorder.

INTRODUCTION

Parkinson's disease is a progressive nervous system disorder that affects movement. Symptoms start gradually, sometimes starting with a barely noticeable tremor in just one hand. Tremors are common, but the disorder also commonly causes stiffness or slowing of movement.
In the early stages of Parkinson's disease, your face may show little or no expression. Your arms may not swing when you walk. Your speech may become soft or slurred. Parkinson's disease symptoms worsen as your condition progresses over time.
Although Parkinson's disease can't be cured, medications might significantly improve your symptoms. Occasionally, your doctor may suggest surgery to regulate certain regions of your brain and improve your symptoms.
How common the condition is that, more than 10 lakh cases per year India. It often requires lab testing Is condition treatable?
Treatments can help manage condition, no known cure. Time taken for recovery. Can last several years or be lifelong.
Parkinson's disease is mainly caused by a loss of nerve cells in substantia nigra of the brain. This leads to a reduction in a chemical called dopamine in the brain. Many different symptoms are associated with Parkinson's disease and the more common symptoms are slowness in movement and muscle stiffness. Parkinson's disease cannot be cured, but medications can help control symptoms. In some later cases, surgery may be advised.

Background and Motivation

Parkinson disease, also called primary parkinsonism, paralysis agitans, or idiopathic parkinsonism, a degenerative neurological disorder that is characterized by the onset of tremor, muscle rigidity, slowness in movement (bradykinesia), and stooped posture (postural instability).
The impact on people as per the common knowledge is, men are been more prone i.e.; about 50%. It was mainly prone for the older people who are 60 years old or more. In these ,10% of patients are diagnosed before they reach their 50.
When someone who is 21-50 years old receives a diagnosis of Parkinson's disease, it is referred to as early onset Parkinson's disease, or young onset Parkinson's disease. While the symptoms of the disease are mostly the same at whatever age it develops, younger people will experience the disease differently due to their unique life circumstances.

It is very important for early detection Parkinson's disease in the affected patients. If diagnosed early the patients can improve their quality of life even if the disease progresses.
The diagnosis is usually based on a careful examination and history of the patient. The doctor will carry out neurological examination and observe the changes in gait, handwriting and facial expressions, difficulty in getting up from a chair, walking etc.
As we gone through many articles, it was said that there was
no cure for it i.e., it has no standard treatment for this problem. As per modern science, Parkinson's treatment varies from patient to patient as per their symptoms. Mainly Parkinson's treatment must include proper medication.

So, as said it shows its eventual progress while reaching to the advanced stage where the patient may loss his life where as if we can detect the problem in the earlier stage so that the patient can get a better life. So, as nearly million people suffering from this problem and losing their life, made us the vision and to know deeper about this project called "Detection of Parkinson's disease".

Problem Statement

The risk factors for Parkinson's disease include:

1.Age -

Young adults rarely experience Parkinson's disease. It ordinarily begins in middle or late life, and the risk increases with age. People usually develop the disease around age 60 or older.

2.Heredity –

Having a close relative with Parkinson's disease increases the chances that you'll develop the disease.

3.Exposure to toxins –

Ongoing exposure to herbicides and pesticides may slightly increase your risk of Parkinson's disease.

So, in order to detect the Parkinson's disease in early stage of the patient body using machine learning. Doctors don't know exactly what causes Parkinson's disease. So, basing on different symptoms the patient is kept under medication. So, basing on those symptoms detecting the disease plays key role.

But the symptoms of the parkinson's disease varies from person to person .so some symptoms may be present in one person and some symptoms may not be. So, here we are taking the vocals as the references with different frequencies and classifying the person's data.

REQUIREMENT ELICITATION AND ANALYSIS

Existing System

In existing system, PD is detected at the secondary stage only (Dopamine deficiency) which leads to medical challenges. Also doctor has to manually examine and suggest medical diagnosis in which the symptoms might vary from person to person so suggesting medicine is also a challenge. Thus the mental disorders are been poorly characterized and have many health complications. PD is generally diagnosed with the following clinical methods.

Proposed System

Here we will be predicting the disease by means of building a system with the

Patient records/data set that consists the features like various frequencies of voice of a particular person that helps to predict the disease yes or no i.e., whether Parkinson's disease is detected or Parkinson's disease not detected. This can be done by a classifier model which helps to tell the person to know about his/her health state .

Feasibility Study

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

It is simply an important outcome of preliminary investigation is the determination that the system request is feasible. This is possible only if it is feasible within limited resources and time.

The different feasibilities that have to be analyzed and the main key consideration involved in the feasibility analysis is /are:

- Operational Feasibility
- Economic Feasibility
- Technical Feasibility

Operational Feasibility: Operational Feasibility deals with the study of prospects of the system to be developed. This system operationally eliminates all the tensions of the admin and helps him in effectively tracking the project progress. This kind of automation will surely reduce the time and energy, which was previously consumed in manual work. Based on the study, the system is proved to be operationally feasible.

Economic Feasibility: Economic Feasibility or Cost-benefit is an assessment of the economic justification for a computer-based project. As hardware was installed from the beginning & for lots of purposes thus the cost on project of hardware is low. Since the system is network based, any number of employees connected to the LAN within that organization can use this tool from any time. The Virtual Private Network is to be developed using the existing resources of the organization. So, the project is economically feasible.

Technical feasibility:

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

System Requirements

The system requirements or software requirements is a listing of what software programs or hardware devices are required to operate the program or game properly. System requirements is a statement that identifies the functionality that is needed by a system in order to satisfy the user's requirements. They are the first and foremost important part of any project, because if the system requirements are not fulfilled, then the project is not complete.

A software requirements is a document that captures complete description about how the system is expected to perform. It is usually signed off at the end of requirements engineering phase.

A software requirement can be of 2 types:

1.Functional Requirements

2.Non-functional Requirements

Functional requirements:

The functional requirements for a system describe what the system should do. Those requirements depend on the type of software being developed, the expected users of the software. These are statement of services the system should provide, how the system should react to particular inputs and how the system should behave in particular situation.

-First, we should import all the necessary datasets.

-We will read the data frame.

-The values in the status are 0 and 1.

-Then fit transform () function is used to fit and transforms data.

SYSTEM REQUIREMENT SPECIFICATION

Functional requirements table

S.no	Requirement	Requirement no.	Essential	Description
1	Loading of data set	RS2	essential	Dataset should be loaded without any errors
2	Data cleaning	RS3	essential	Data should be cleaned properly
3	Feature Selection	RS4	essential	Feature selection is performed
4	Splitting of dataset into train and test data	RS5	essential	According to the given ratio the dataset should be split into train and test data
5	Random forest algorithm is called for given data set	RS6	essential	the model applied should work without any errors
6	prediction	RS7	essential	Prediction should done without any error

SYSTEM DESIGN

System Design is a solution to how to approach to the creation of a system. System Design is the process of designing the architecture, components, and interfaces for a system so that it meets the end- user requirements. System design is the process of defining the architecture, components, modules, interfaces and data for a system to satisfy specified requirements. System design could be seen as the application of the system theory to product development. There is some overlap with the disciplines of the system analysis, system architecture and system engineering. System modelling is the interdisciplinary study of the use of mobile to conceptualize and construct system in business and IT development. This important phase provides the understanding and procedural details necessary for implementing the system recommended in the feasibility study. The design step produces a data design, an architectural design and a procedural design. The data design transforms the information domain model created during analysis in to the data structures that will be required to implement the software. The architectural design defines the relationship among major structural components into a procedural description of the software.

Source code generated and testing is conducted to integrate and validate the software. From a project management point of view, software design is conducted in two steps.

Preliminary design is connected with the transformation of requirements into data and software architecture.

Detailed design focuses on refinements to the architectural representation that leads the detailed data structure and algorithmic representations of software.

Reliability in System Design

A system is Reliable when it can meet the end-user requirement. When designing it, we should have planned to implement a set of features and services in your system. If the system can serve all those features without wearing out then your System can be considered to be Reliable.

A Fault Tolerant system can be one that can continue to be functioning reliably even in the presence of faults. Faults are the errors that arise in a particular component of the system. An occurrence of fault doesn't guarantee Failure of the System. Failure is the state when the system is not able to perform as expected. It is no longer able to provide certain services to the end-users.

Availability in System Design

Availability is a characteristic of a System which aims to ensure an agreed level of Operational Performance, also known as uptime. It is essential for a system to ensure high availability in order to serve the user's requests. The extent of Availability varies from system to system.

"Suppose when we are designing a Social Media Application then high availability is not much of a need. A delay of a few seconds can be tolerated. Getting to view the post on Instagram with a delay of 5 to 10 seconds will not be much of an issue. But if you are designing a system for hospitals, Data center, or Banking, then you should ensure that your system is highly available. Because a delay in the service can lead to a huge loss."

So now as we are designing a system regarding hospitals, we should make our system is highly available.

There are various principles you should follow in order to ensure the availability of your system:

•System should not have a Single Point of Failure. Basically, your system should not be dependent on a single service in order to process all of its requests.

•Because when that service fails then your entire system can be end up becoming unavailable.

•Detecting the Failure and resolving it at that point.

Scalability in System

Design Scalability refers to the ability of the System to cope up with the increasing load. While designing the system you should keep in mind the load experienced by it. It's said that if you have to design a system for load X then you should plan to design it for 10X and test it for 100X. There can be a situation where your system can experience an increasing load.

Swimlane Activity diagram:

The Activity diagrams in Object Oriented design are just like the flow carts that show the sequence of steps that make up a complex process, such as an algorithm or workflow. Activity diagram are most useful during the initial stages of the design phase.

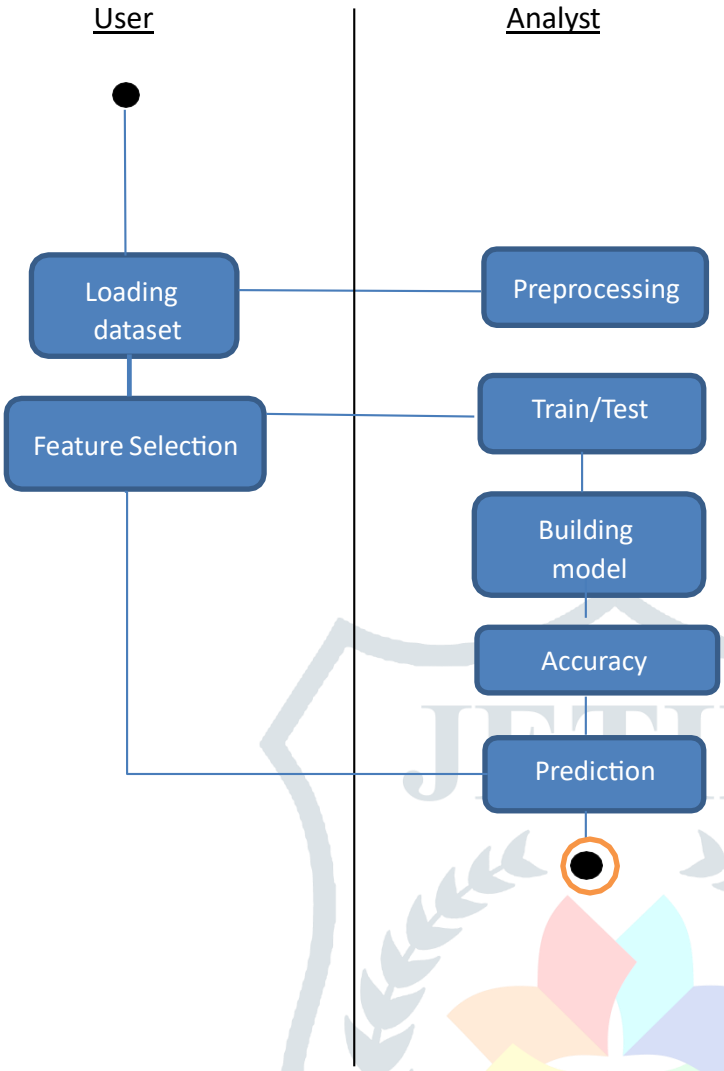
These can be both sequential and in parallel. They describe the objects used, consumed or produced by an activity and the relationship between the different activities.

Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part. It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single. Before drawing an activity diagram, we must have a clear understanding about the elements used in the activity diagram. The main element of an activity diagram is the activity itself. Activity is a function performed by the system. After identifying the activities, we need to understand how they are associated with constraints and conditions.

Swimlanes are used to show which activities are performed by which organization in the activity diagram. The lanes are boundaries are drawn and the activities of a particular organization are drawn in the same lane as that of the organization.

Swim lanes have to be ordered in a Logical Manner. It is suggested to have less than five swim lanes in an activity diagram. Swim lanes are good in that they combine the activity diagram's depiction of logic with the interaction diagram's depiction of responsibility.

Swimlane diagram (use case +activity)



State Chart Diagram:

A State chart diagram describes a state machine. State machine can be defined as a machine which defines different states of an object and these states are controlled by external or internal events. As the State chart diagram defines the states, it is used to model the lifetime of an object. State chart diagram describes the flow of control from one state to another state. States are defined as a condition in which an object exists and it changes when some event is triggered. The most important purpose of a state chart diagram is to model the lifetime of an object from creation to termination

Dataset :

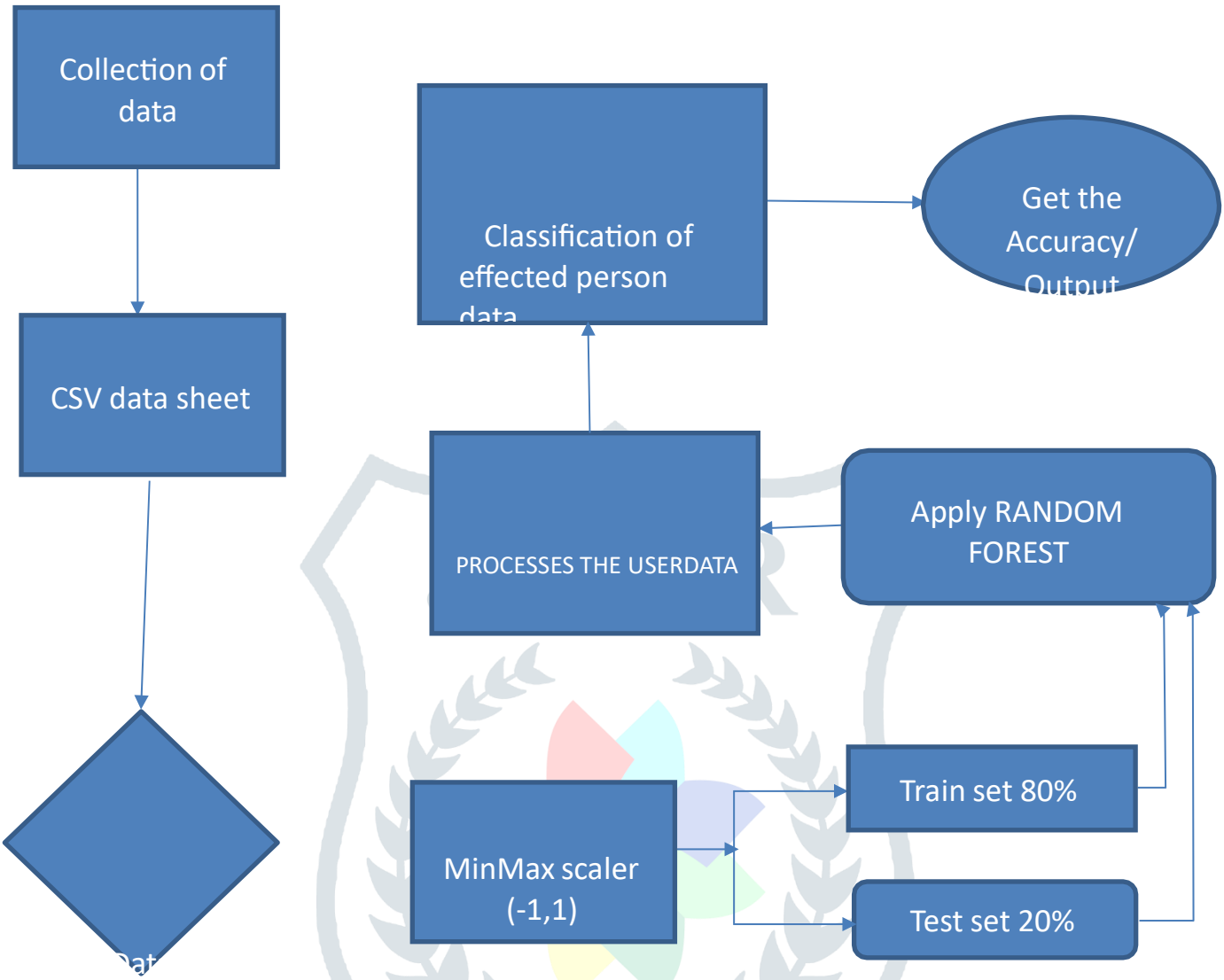
The most commonly marked effect is on speech, including (dysarthria) difficulty articulating sounds, (hypophonia) lowered volume, and (monotone) reduced pitch range. The dataset mainly contains 195 entries that containing with the information of the individual's vocal records.

And these vocal records are specified in various forms in 24 columns. It contains average vocal fundamental frequency, maximum vocal fundamental frequency, minimum vocal fundamental frequency, etc. The dataset consists the information of both Parkinson disease effected person and healthy person.

It follows the flow of process as following

•Collection of data
•Loading csv data sheet
•Data pre-processing
•Splitting the data set into train and tests sets
•Applying Random Forest Algorithm
•Processes the user data
•Classification of effected person data
•the accuracy of model
•Predicts that the person is with the Parkin

The following diagram depicts the entire process the prediction system:



is the execution or practice of a plan, a method or any design, idea, model, specification, standard or policy for doing something. As such, implementation is the action that must follow any preliminary thinking for something to actually happen.

To provide as much flexibility as possible, the display application is implemented as multiple nearly- independent modules. Each module is responsible for the display of a particular kind of data and is implemented.

•analyzing requirements
•installation
•configuration
•customization
•testing
•running
•systems integrations
•user training
•delivery
•making changes

Software Environment

A software development environment (SDE) is the collection of hardware and software tools a system developer uses to build software systems. When you are developing software, you probably don't want your users to see every messy part of your application creation process

There 4 different environments in a software development team are shown below: Development environment.

Testing environment. Staging environment. Production environment.

Software Technologies

The software technology used in this project is python.

Python is the fastest growing programming language.It supports multiple programming paradigms,including structured,object-oriented and functional programming.And it is dynamically- typed and garbage collected.

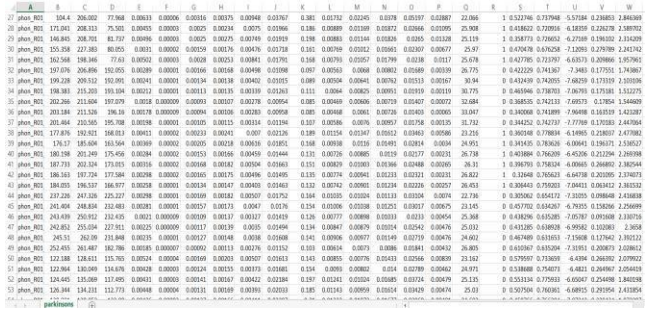
It consistently ranks as one of the most popular programming languages.It can be also used on a server to create web applications .
It has a huge number of libraries and frameworks.
Python frameworks are no different; they are a collection of modules and packages.
These frameworks automate common processes and implementation. For instance, developers can focus on application logic rather than dealing with routinary processes.
The python libraries used are:
•numpy
•pandas
•matplotlib
•sklearn
•seaborn

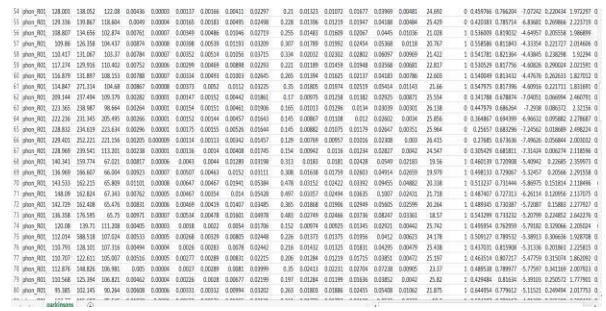
Numpy-
The name "Numpy" stands for "Numerical Python". It is the commonly used library. It is a popular machine learning library that supports large matrices and multi-dimensional data. It consists of in-built mathematical functions for easy computations. Even libraries like TensorFlow use Numpy internally to perform several operations on tensors. Array Interface is one of the key features of this library
Pandas-
pandas is a software library written for the Python programming language for data manipulation and analysis. When we have to work on Tabular data, we prefer the pandas module. The powerful tools of pandas are Data frame and Series. Pandas has a better performance when a number of rows is 500K or more.
Matplotlib-
matplotlib() is a library function that is responsible for plotting numerical data. And that's why it is used in data analysis. It is also an open-source library and plots high-defined figures like pie charts, histograms, scatterplots, graphs, etc.
Scikit-learn-
We use this as Scikit-learn (). It is a famous Python library to work with complex data. Scikit-learn is an open-source library that supports machine learning. It supports variously supervised and unsupervised algorithms like linear regression, classification, clustering, etc. This library works in association with NumPy and SciPy.
Seaborn-
Seaborn is a library that uses Matplotlib underneath to plot graphs. It will be used to visualize random distributions. It is used for data visualization and exploratory data analysis. Seaborn works easily with
Random Forest Algorithm-
Random forest is a supervised learning algorithm. It can be used for both Classification and Regression problems. It uses an ensemble learning method known as 'bagging' (Bootstrap Aggregation) which is a process of combining multiple classifiers to solve a complex problem.
Random forest creates various random subsets of the given dataset and passes them to different numbers of decision trees and takes the prediction from each tree. Based on the majority votes of prediction, random forest takes the average to predict the output and also to increase the accuracy and overall result. As there are greater numbers of trees in random forests, it prevents the problem of overfitting.
Random forest searches for the most important feature while splitting a node which helps in building a better model.
Python Implementation of Random Forest Algorithm
Here we will implement the Random Forest Algorithm tree using Python. For this, we will use the same dataset "user_data.csv", which we have used in classification models. By using the same dataset, we can compare the Random Forest classifier with other classification models such as Decision tree Classifier, KNN, SVM, Logistic Regression, etc.
Implementation Steps are given below:
•Data Pre-processing step
•Fitting the Random forest algorithm to the Training set
•Predicting the test resultTest accuracy of the result. (Creation of Confusion matrix)
•Visualizing the test set result.

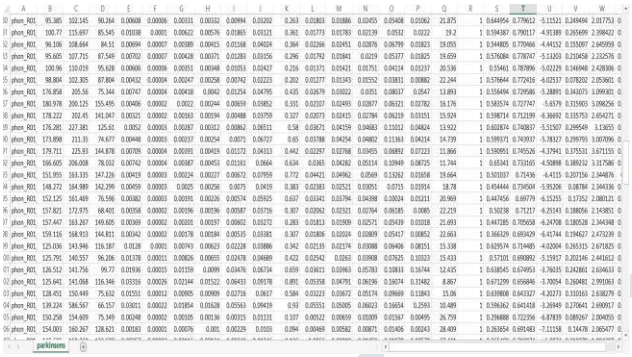
csv file

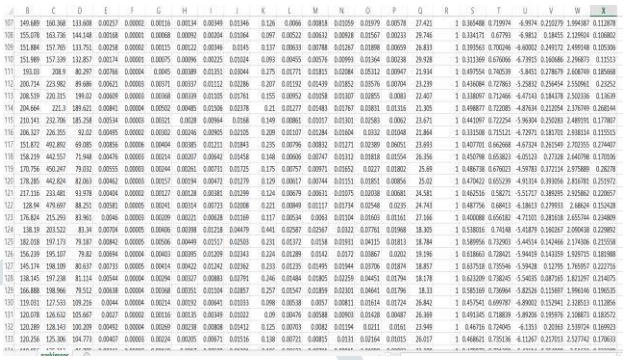
	name	MDVP:Fo(I	MDVP:Fhi(	MDVP:Flo(	MDVP:Jitte	MDVP:Jitte	MDVP:RAF	MDVP:PPC	Jitter:DDP	MDVP:Shir	MDVP:Shir	Shimmer:A	Shimmer:A	MDVP:APC	Shimmer:D	NHR	HNR	status	RPDE	DFA	spread1	spread2	D2	
2	phon_R01	119.992	157.302	74.997	0.00784	0.00007	0.0037	0.00554	0.01109	0.04374	0.426	0.02182	0.0313	0.02971	0.06545	0.02211	21.033	1	0.414783	0.815285	-4.81303	0.266482	2.301442	
3	phon_R01	122.4	148.65	113.819	0.00968	0.00008	0.00465	0.00696	0.01394	0.06134	0.626	0.03134	0.04518	0.04368	0.09403	0.01929	19.085	1	0.458359	0.819521	-4.07519	0.33559	2.486855	
4	phon_R01	116.682	131.111	111.555	0.0105	0.00009	0.00544	0.00781	0.01633	0.05233	0.482	0.02757	0.03858	0.0359	0.0827	0.01309	20.651	1	0.429895	0.825288	-4.44318	0.311173	2.342259	
5	phon_R01	116.676	137.871	111.366	0.00997	0.00009	0.00502	0.00698	0.01505	0.05492	0.517	0.02924	0.04005	0.03772	0.08771	0.01353	20.644	1	0.434969	0.819235	-4.1175	0.334147	2.405554	
6	phon_R01	116.014	141.781	110.655	0.01284	0.00011	0.00655	0.00908	0.01966	0.06425	0.584	0.0349	0.04825	0.04465	0.1047	0.01767	19.649	1	0.417356	0.823484	-3.74779	0.234513	2.33218	
7	phon_R01	120.552	131.162	113.787	0.00968	0.00008	0.00463	0.0075	0.01388	0.04701	0.456	0.02328	0.03526	0.03243	0.06985	0.01222	21.378	1	0.415564	0.825069	-4.24287	0.299111	2.18756	
8	phon_R01	120.267	137.244	114.82	0.00333	0.00003	0.00155	0.00202	0.00466	0.01608	0.14	0.00779	0.00937	0.01351	0.02337	0.00607	24.886	1	0.59604	0.764112	-5.63432	0.257682	1.854785	
9	phon_R01	107.332	113.84	104.315	0.0029	0.00003	0.00144	0.00182	0.00431	0.01567	0.134	0.00829	0.00946	0.01256	0.02487	0.00344	26.892	1	0.63742	0.763262	-6.1676	0.183721	2.064693	
10	phon_R01	95.73	132.068	91.754	0.00551	0.00006	0.00293	0.00332	0.0088	0.02093	0.191	0.01073	0.01277	0.01717	0.03218	0.0107	21.812	1	0.615551	0.773587	-5.49868	0.327769	2.322511	
11	phon_R01	95.056	120.103	91.226	0.00532	0.00006	0.00268	0.00332	0.00803	0.02838	0.255	0.01441	0.01725	0.02444	0.04324	0.01022	21.862	1	0.547037	0.798463	-5.01188	0.325996	2.432792	
12	phon_R01	88.333	112.24	84.072	0.00505	0.00006	0.00254	0.0033	0.00763	0.02143	0.197	0.01079	0.01342	0.01892	0.03237	0.01166	21.118	1	0.611137	0.776156	-5.24977	0.391002	2.407313	
13	phon_R01	91.904	115.871	86.292	0.0054	0.00006	0.00281	0.00336	0.00844	0.02752	0.249	0.01424	0.01641	0.02214	0.04272	0.01141	21.414	1	0.58339	0.79252	-4.96023	0.363566	2.642476	
14	phon_R01	136.926	159.866	131.276	0.00293	0.00002	0.00118	0.00153	0.00355	0.01259	0.112	0.00656	0.00717	0.0114	0.01968	0.00581	25.703	1	0.4606	0.646846	-6.54715	0.152813	2.041277	
15	phon_R01	139.173	179.139	76.556	0.0039	0.00003	0.00165	0.00208	0.00496	0.01642	0.154	0.00728	0.00932	0.01797	0.02184	0.01041	24.889	1	0.430166	0.665833	-5.66022	0.254989	2.519422	
16	phon_R01	152.845	163.305	75.836	0.00294	0.00002	0.00121	0.00149	0.00364	0.01828	0.158	0.01064	0.00972	0.01246	0.03191	0.00609	24.922	1	0.474791	0.654027	-6.1051	0.203653	2.125618	
17	phon_R01	142.167	217.455	83.159	0.00369	0.00003	0.00157	0.00203	0.00471	0.01503	0.126	0.00772	0.00888	0.01359	0.02316	0.00839	25.175	1	0.565924	0.658245	-5.34012	0.210185	2.205546	
18	phon_R01	144.188	349.259	82.764	0.00544	0.00004	0.00211	0.00292	0.00632	0.02047	0.192	0.00969	0.012	0.02074	0.02908	0.01859	22.333	1	0.56738	0.644692	-5.44004	0.239764	2.264501	
19	phon_R01	168.778	232.181	75.603	0.00718	0.00004	0.00284	0.00387	0.00853	0.03327	0.348	0.01441	0.01893	0.0343	0.04322	0.02919	20.376	1	0.631099	0.605417	-2.93107	0.434326	3.007463	
20	phon_R01	153.046	175.829	68.623	0.00742	0.00005	0.00364	0.00432	0.01092	0.05517	0.542	0.02471	0.03572	0.05767	0.07413	0.0316	17.28	1	0.665318	0.719467	-3.94908	0.35787	3.10901	
21	phon_R01	156.405	189.398	142.822	0.00768	0.00005	0.00372	0.00399	0.01116	0.03995	0.348	0.01721	0.02374	0.0431	0.05164	0.03365	17.153	1	0.649554	0.68608	-4.55447	0.340176	2.856676	
22	phon_R01	153.848	165.738	65.782	0.0084	0.00005	0.00428	0.0045	0.01285	0.0381	0.328	0.01667	0.02383	0.04055	0.05	0.03871	17.536	1	0.660125	0.704087	-4.09544	0.262564	2.73971	
23	phon_R01	153.88	172.86	78.128	0.0048	0.00003	0.00232	0.00267	0.00696	0.04137	0.37	0.02021	0.02591	0.04525	0.06062	0.01849	19.493	1	0.629017	0.698951	-5.18696	0.237622	2.557536	
24	phon_R01	167.93	193.221	79.068	0.00442	0.00003	0.0022	0.00247	0.00661	0.04351	0.377	0.02228	0.0254	0.04246	0.06685	0.0128	22.468	1	0.61906	0.679834	-4.33096	0.262384	2.916777	
25	phon_R01	173.917	192.735	86.18	0.00476	0.00003	0.00221	0.00258	0.00663	0.04192	0.364	0.02187	0.0247	0.03772	0.06562	0.0184	20.422	1	0.537264	0.686894	-5.24878	0.210279	2.547508	
26	phon_R01	163.656	200.841	76.779	0.00742	0.00005	0.0038	0.0039	0.0114	0.01659	0.164	0.00738	0.00948	0.01497	0.02214	0.01778	23.831	1	0.397937	0.732479	-5.55745	0.22089	2.692176	
27	phon_R01	104.4	206.002	77.968	0.00633	0.00006	0.00316	0.00375	0.00948	0.03767	0.381	0.01732	0.02245	0.0378	0.05197	0.02887	22.066	1	0.522746	0.737948	-5.57184	0.236853	2.846369	

parkinsons



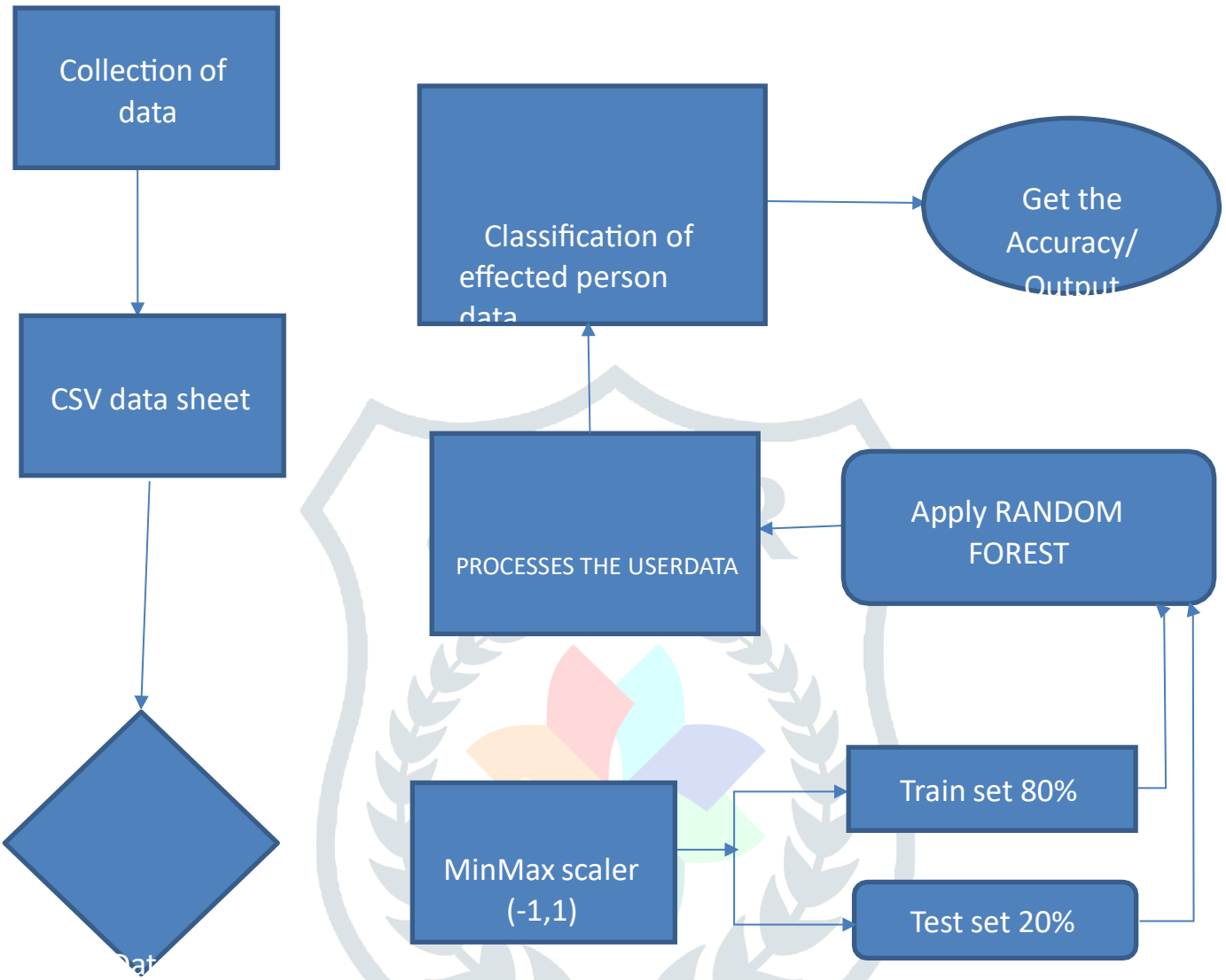






	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
170	197.569	217.627	90.794	0.00803	0.00004	0.0049	0.00448	0.0147	0.02177	0.189	0.01279	0.01272	0.01439	0.03836	0.01337	19.269	0	0.372222	0.725216	-5.73678	0.164529	2.88245	0.202879
171	240.301	245.135	219.783	0.00517	0.00002	0.00316	0.00283	0.00949	0.02018	0.212	0.01176	0.01191	0.01344	0.03529	0.00965	21.02	0	0.371837	0.646167	-7.1697	0.073298	2.266432	0.100881
172	244.99	272.21	239.17	0.00451	0.00002	0.00279	0.00237	0.00837	0.01897	0.181	0.01084	0.01121	0.01255	0.03253	0.01049	21.528	0	0.522812	0.646818	-7.3045	0.171088	2.095237	0.09622
173	112.547	133.374	105.715	0.00355	0.00003	0.00166	0.0019	0.00499	0.01358	0.129	0.00664	0.00786	0.0114	0.01992	0.00435	26.436	0	0.413295	0.7567	-6.32353	0.218885	2.193412	0.160376
174	110.739	113.597	100.139	0.00356	0.00003	0.0017	0.002	0.0051	0.01484	0.133	0.00754	0.0095	0.01285	0.02261	0.0043	26.55	0	0.36909	0.776158	-6.08557	0.192375	1.889002	0.174152
175	113.715	116.443	96.913	0.00349	0.00003	0.00171	0.00203	0.00514	0.01472	0.133	0.00748	0.00905	0.01148	0.02245	0.00478	26.547	0	0.380253	0.7667	-5.9435	0.19215	1.852542	0.179677
176	117.004	144.466	99.923	0.00353	0.00003	0.00176	0.00218	0.00528	0.01657	0.145	0.00881	0.01062	0.01318	0.02643	0.0059	25.445	0	0.387482	0.756482	-6.01256	0.229298	1.872946	0.163118
177	115.38	123.109	108.634	0.00332	0.00003	0.0016	0.00199	0.0048	0.01503	0.137	0.00812	0.00933	0.01133	0.02436	0.00401	26.005	0	0.405991	0.761255	-5.96678	0.197938	1.974857	0.184067
178	116.388	129.038	108.97	0.00346	0.00003	0.00169	0.00213	0.00507	0.01725	0.155	0.00874	0.01021	0.01331	0.02623	0.00415	26.143	0	0.361232	0.763242	-6.01689	0.109256	2.004719	0.174429
179	151.737	190.204	129.859	0.00314	0.00002	0.00135	0.00162	0.00406	0.01469	0.132	0.00728	0.00886	0.0123	0.02184	0.0057	24.151	1	0.39661	0.745957	-6.48682	0.197919	2.449763	0.132703
180	148.79	158.359	138.99	0.00309	0.00002	0.00152	0.00186	0.00456	0.01574	0.142	0.00839	0.00956	0.01309	0.02518	0.00488	24.412	1	0.402591	0.762508	-6.31199	0.182459	2.251553	0.160306
181	148.143	155.982	135.041	0.00392	0.00003	0.00204	0.00231	0.00612	0.0145	0.131	0.00725	0.00876	0.01263	0.02175	0.0054	23.683	1	0.398499	0.778349	-5.71121	0.240875	2.845109	0.19273
182	150.44	163.441	144.736	0.00396	0.00003	0.00206	0.00233	0.00619	0.02551	0.237	0.01321	0.01574	0.02148	0.03964	0.00611	23.133	1	0.352396	0.75932	-6.26145	0.183218	2.264226	0.144105
183	148.462	161.078	141.998	0.00397	0.00003	0.00202	0.00235	0.00605	0.01831	0.163	0.0095	0.01103	0.01559	0.02849	0.00639	22.866	1	0.408598	0.768845	-5.70405	0.216204	2.679185	0.19771
184	149.818	163.417	144.786	0.00336	0.00002	0.00174	0.00198	0.00521	0.02145	0.198	0.01155	0.01341	0.01666	0.03464	0.00595	23.008	1	0.329577	0.75718	-6.27717	0.109397	2.209021	0.156368
185	117.226	123.925	106.656	0.00417	0.00004	0.00186	0.0027	0.00558	0.01909	0.171	0.00864	0.01223	0.01949	0.02592	0.00955	23.079	0	0.603515	0.669565	-5.61907	0.191576	2.027228	0.215724
186	116.848	217.552	99.503	0.00531	0.00005	0.0026	0.00346	0.0078	0.01795	0.163	0.0081	0.01144	0.01756	0.02429	0.01179	22.085	0	0.663842	0.656516	-5.19886	0.206768	2.120412	0.252404
187	116.286	177.291	96.983	0.00314	0.00003	0.00134	0.00192	0.00403	0.01564	0.136	0.00667	0.0099	0.01691	0.02001	0.00737	24.199	0	0.598515	0.654331	-5.59258	0.133917	2.058658	0.214346
188	116.556	592.03	86.228	0.00496	0.00004	0.00254	0.00263	0.00762	0.0166	0.154	0.0082	0.00972	0.01491	0.0246	0.01397	23.958	0	0.566424	0.667654	-6.43112	0.15331	2.161936	0.120605
189	116.342	581.289	94.246	0.00267	0.00002	0.00115	0.00148	0.00345	0.013	0.117	0.00631	0.00789	0.01144	0.01892	0.0068	25.023	0	0.528485	0.663884	-6.35902	0.116636	2.152083	0.138868
190	114.563	119.167	86.647	0.00327	0.00003	0.00146	0.00184	0.00439	0.01185	0.106	0.00557	0.00721	0.01095	0.01672	0.00703	24.775	0	0.555303	0.659132	-6.71022	0.149694	1.91399	0.121777
191	201.774	262.707	78.228	0.00694	0.00003	0.00412	0.00396	0.01235	0.02574	0.255	0.01454	0.01582	0.01758	0.04363	0.04441	19.368	0	0.508479	0.683761	-6.93447	0.15989	2.316346	0.112838
192	174.188	230.978	94.261	0.00459	0.00003	0.00263	0.00259	0.0079	0.04087	0.405	0.02336	0.02498	0.02745	0.07008	0.02764	19.517	0	0.448439	0.657899	-6.53859	0.121952	2.657476	0.13305
193	209.516	253.017	89.488	0.00564	0.00003	0.00331	0.00292	0.00994	0.02751	0.263	0.01604	0.01657	0.01879	0.04812	0.0181	19.147	0	0.431674	0.683244	-6.19533	0.129303	2.784312	0.168895
194	174.688	240.005	74.287	0.0136	0.00008	0.00624	0.00564	0.01873	0.02308	0.256	0.01268	0.01365	0.01667	0.03804	0.10715	17.883	0	0.407567	0.655683	-6.7872	0.158453	2.679772	0.131728
195	198.764	396.961	74.904	0.0074	0.00004	0.0037	0.0039	0.01109	0.02296	0.241	0.01265	0.01321	0.01588	0.03794	0.07223	19.02	0	0.451221	0.643956	-6.74458	0.207454	2.138608	0.123306
196	214.289	260.277	77.973	0.00567	0.00003	0.00295	0.00317	0.00885	0.01884	0.19	0.01026	0.01161	0.01373	0.03078	0.04398	21.209	0	0.462803	0.664357	-5.72406	0.190667	2.555477	0.148569

The following diagram depicts the entire process the prediction system:



is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. In simple words, testing is executing a system in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements.

The main purpose of testing is to discover errors and it is a process to check the functionality of components. It ensures that software systems meets its requirements and user expectations without fault .

Software Testing is an important element of quality assurance and represents ultimate view of specification and design.

An early start to testing reduces the cost and time to rework and produce error-free software that is delivered to the client. However in Software Development Life Cycle (SDLC), testing can be started from the Requirements Gathering phase and continued till the deployment of the software.

It also depends on the development model that is being used. For example, in the Waterfall model, formal testing is conducted in the testing phase; but in the incremental

Testing

is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. In simple words, testing is executing a system in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements.

The main purpose of testing is to discover errors and it is a process to check the functionality of components. It ensures that software systems meets its requirements and user expectations without fault .

Software Testing is an important element of quality assurance and represents ultimate view of specification and design.

Testing can also be defined as –

A process of analyzing a software item to detect the differences between existing and required conditions (that is defects/errors/bugs) and to evaluate the features of the software item.

When to Start Testing?

An early start to testing reduces the cost and time to rework and produce error-free software that is delivered to the client. However in Software Development Life Cycle (SDLC), testing can be started from the Requirements Gathering phase and continued till the deployment of the software.

It also depends on the development model that is being used. For example, in the Waterfall model, formal testing is conducted in the testing phase; but in the incremental model, testing is performed at the end of every increment/iteration and the whole application is tested at the end.

Testing is done in different forms at every phase of SDLC −

During the requirement gathering phase, the analysis and verification of requirements are also considered as testing.

Reviewing the design in the design phase with the intent to improve the design is also considered as testing.

Testing performed by a developer on completion of the code is also categorized as testing.

When to Stop Testing?

It is difficult to determine when to stop testing, as testing is a never-ending process and no one can claim that a software is 100% tested. The following aspects are to be considered for stopping the testing process −

Testing Deadlines

Completion of test case execution

Completion of functional and code coverage to a certain point

White Box Testing

White box testing is a software testing method in which the internal structure/design/implementation of the item being tested is known to the tester. Implementation and impact of the code are tested.

It is a way of testing the software in which the tester has knowledge about the internal structure or the code or the program of the software.

•code implementation is necessary for white box testing.

•It is mostly done by software developers.

•Knowledge of implementation is required.

•It is the inner or the internal software testing.

•It is the Structural test if the software.

•This type of testing is software is started after detail design document.

Test case table-1

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
1	TC1	Loading of dataset	RS2	essential	Dataset is downloaded and then loaded	The dataset should be loaded without any errors	The dataset is loaded without any errors	SUCCESS

Test case table-2

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
2	TC2	Data Cleaning	RS3	essential	Dataset is checked for null values, if any null values are found they should be handled	Data should be cleaned and should and should contain no null values	No null values are found	SUCCESS

Test case table-3

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
3	TC3	Feature Selection	RS4	essential	Feature selection is done using heatmap	Heatmap shows whether all columns are relevant to the target or not	Heatmap is generated and all the features are found to be significant	SUCCESS

Test case table-4

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
4	TC4	Splitting of the dataset into train and test data	RS5	essential	Dataset is split into train and test data in the ratio 8:2	According to the given ratio the dataset should be split into train and test data	The given dataset is split into train and test data in 8:2 ratio.	SUCCESS

Test case table-5

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
5	TC5	Random forest algorithm is called for given data set	RS6	essential	Random forest algorithm is applied	The model applied should work without any errors	Result is classified (yes/no)	SUCCESS

Test case table-6

S.no	Test Case no.	Requirements (req)	Req no.	Essential	Description	Expected Output	Actual Output	Result
6	TC6	Evaluation of model	RS7	essential	Accuracy, precision are measured	evaluation properties are measured properly	All the properties are measured and compared	SUCCESS

CONCLUSION

Parkinson's disease affects the CNS of the brain and has yet no treatment unless it's detected early. Late detection leads to no treatment and loss of life. Thus its early detection is significant. For early detection of the disease, we utilized machine learning algorithms.

This Project entitles "Machine learning based parkinson's disease prediction system" which attempts to predict the health status of the person (i.e, whether the person is suffering from parkinson's disease or not).In this Machine learning project, we developed a module using the Random Forest Classifier of the sklearn module of python to detect if an individual has Parkinson's Disease or not, We got the machine learning model with 97.43% accuracy ,which is good as our dataset contains less records. Where the existed system has the accuracy of 94.91% with the same classifier that is used in the proposed system.

FUTURE SCOPE

No cure for Parkinson's Disease exists today, but research is ongoing and medications or surgery can often provide substantial improvement with motor symptoms. It would be better if there is a website that helps the patients those are suffering from Parkinson's

Disease by providing a proper health care remedies, not only detecting through the limited features but also including the symptoms of the disease. But as these symptoms become more obvious, people may have difficulty walking, talking, or completing other simple tasks. Not everyone with one or more of these symptoms has PD, as the symptoms appear in other diseases as well.

REFRENCES

[1] A. Tzanetakis, G. and Cook, P. "Musical genre classification of audio signal", IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 3, pp. 293-302, July 2002.

[2] Holzapfel, A. and Stylianou Y. "Musical genre classification using nonnegative matrix factorization-based features", IEEE Transactions on Audio, Speech, and Language Processing, Vol. 16, No. 2, pp. 424-434, 2008.

[3] Benetos, E. and Kotropoulos C. "A tensor-based approach forautomatic music genre classification", Proceedings of the European Signal Processing Conference, Lausanne, Switzerland, 2008

[4] Lidy, T., Rauber, A., Pertusa, A. and Inesta, J. "Combining audio and symbolic descriptors for music classification from audio", Music Information Retrieval Information Exchange (MIREX), 2007. [5] Bergstra, J., Casagrande, N., Erhan, D., Eck, D. and Kegl B. "Aggregate features and AdaBoost for music classification", Machine Learning, Vol. 65, No. 2-3, pp. 473-484, 2006.

[6] A. Karatana and O. Yildiz, "Music genre classification with machine learning techniques," in Signal Processing and Communications Applications Conference (SIU), 2017 25th, IEEE, 2017, pp. 1–4.

[7] Han, J., Kamber, M., & Pei, J. (2011). Data mining: concepts and techniques: concepts and techniques. Elsevier.

[8] Witten, I. H., & Frank, E. (2005). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

[9] Cortez P., In press. RMiner: Data Mining with Neural Networks and Support Vector Machines using R. In R. Rajesh (Ed.), Introduction to Advanced Scientific Softwares and Toolboxes.

[10] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5– 32, Oct. 2001.

[11] G. Gwardys and D. Grzywczak, "Deep image features in music information retrieval," Intl Journal of Electronics and Telecomunications, vol. 60, no. 4, pp. 321–361, 2014.

[12]https://librosa.github.io/librosa/(visitedon15/06/2018)

[13] https://keras.io (visited on 15/06/2018)

[14] S. Dieleman. (2014). Recommending music on spotify with deep learning,[Online].Available: http://benanne.github.io/2014/08/05/spotifycnns.html (visited on 03/31/2018

[15] Ahmet Elbir, Hamza Osman Ilhan, Gorkem Serbes, Nizamettin Aydin. "Short Time Fourier Transform based music Genre classification" , 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), 2018

BIBILIOGRAPHY

V.Neelima Devi Project Guide working as head of department in Master of computer application (MCA) Sanketika Vidya Parishad Engineering college, Visakhapatnam, Andhrapradesh. with 6 years of experience in master of computer applications(MCA),Accredited by NAAC. with her area of interests in C, Computer Organization, Software Engineering IOT.

Rampi Sireesha is studying her 2nd year, Master of Computer Applications in Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, accredited by NAAC. With her interest in python, machine learning and as a part of academic project, she choose Image processing using python. As a result of a desire to comprehend the flaws in conventional reporting and to preserve timely and high-quality report output in Face mask recognition. A completely developed project along with code has been submitted for Andhra University as an Academic Project. In completion of her MCA.