KIET GROUP OF INSTITUTIONS
Connecting Life with Learning
www.kiet.edu
Delhi-NCR, Ghaziabad

PARKINSONS DISEASE PREDICTION USING MACHINE LEARNING

Group ID: PCSE25-70

Team Leader : Varun gupta

Group Member: Varun gupta	**Semester :** 8	**Department:** CSE
Group Member: Sumit pal	**Semester :** 8	**Department:** CSE
Group Member: Yash jain	**Semester :** 8	**Department:** CSE

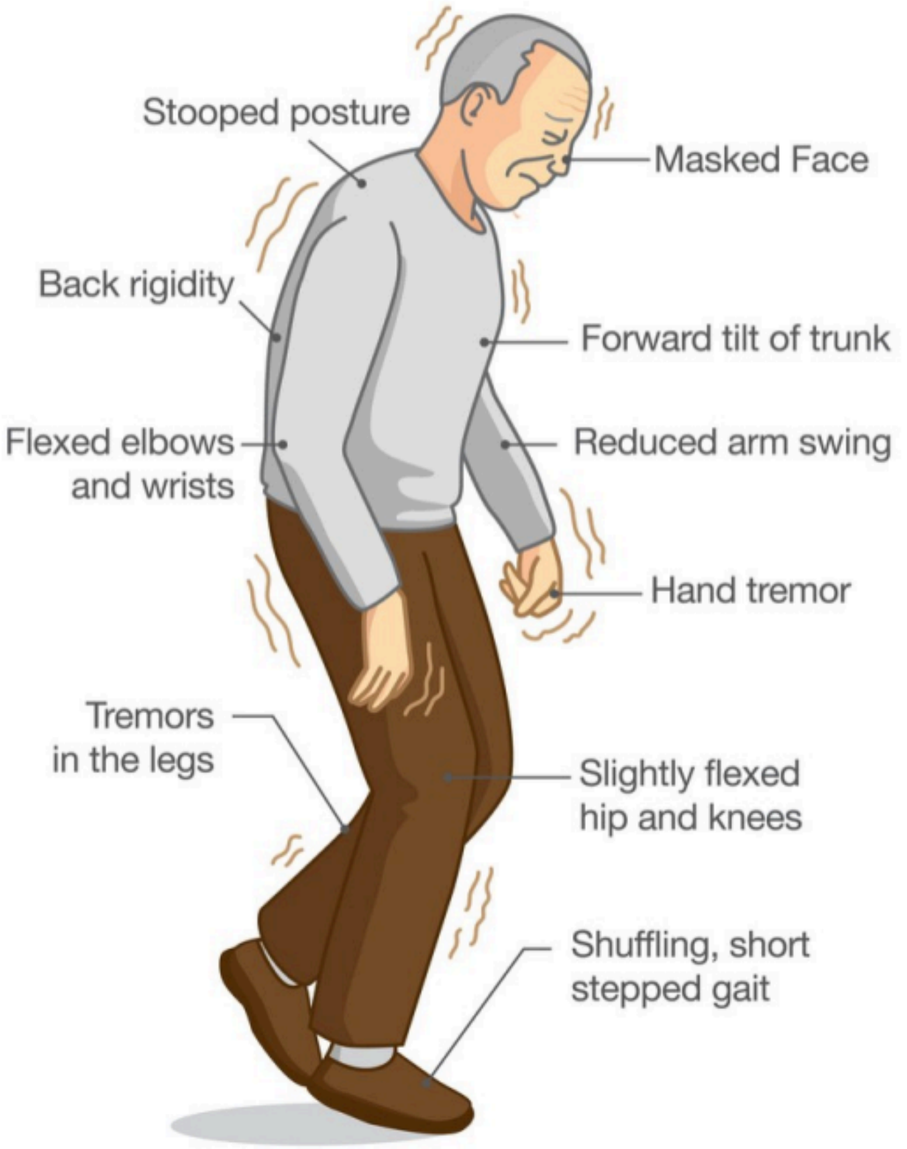
Name of Guide : Swati Sharma

Date of Presentation : 24/05/2025

Introduction

Parkinson’s disease is a movement disorder that affects the central nervous system. The symptoms start gradually because of low dopamine levels in the brain. Dopamine is a chemical and is a neurotransmitter responsible for sending signals from the body to the brain. Reduction in the neurons responsible for production of dopamine leads to lowering of dopamine levels resulting in reduced coordination between the brain and the body. Till date there's not any cure for this disease and with the advancement of technology, it is necessary to introduce a quick and reasonable tool to predict the disease.

Parkinson's Disease Symptoms



Symptoms associated with PD depending on stage Machine learning can be used as a powerful tool for the prediction of various kinds of disease. It can help to predict the presence of PD more accurately and cost effectively. One of the main symptoms of PD is loss of speech. The resultant problems may include monotonic speech, slur words, mumbling, breathy or hoarse speech etc. Keeping the points in mind, we decided to go with a voice dataset containing 25 features. In order to select the best features, Pearson method of correlation is used.

In this project, we present a prediction system for PD using the XGBoost model and the Pearson method of correlation for feature selection to predict the presence of Parkinson‟s Disease.

The algorithms used in the analysis for choosing the optimistic algorithms are:

1. SVM
2. Random Forest
3. XGBoost

	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP	MDVP:Shimmer
MDVP:Fo(Hz)	1	0.400985	0.596546	-0.118003	-0.382027	-0.0761938	-0.112165	-0.0762127	-0.0983737
MDVP:Fhi(Hz)	0.400985	1	0.0849513	0.102086	-0.0291983	0.0971766	0.0911262	0.0971499	0.00228123
MDVP:Flo(Hz)	0.596546	0.0849513	1	-0.139919	-0.277815	-0.100519	-0.0958284	-0.100488	-0.144543
MDVP:Jitter(%)	-0.118003	0.102086	-0.139919	1	0.935714	0.990276	0.974256	0.990276	0.769063
MDVP:Jitter(Abs)	-0.382027	-0.0291983	-0.277815	0.935714	1	0.922911	0.897778	0.922913	0.703322
MDVP:RAP	-0.0761938	0.0971766	-0.100519	0.990276	0.922911	1	0.957317	1	0.759581
MDVP:PPQ	-0.112165	0.0911262	-0.0958284	0.974256	0.897778	0.957317	1	0.957319	0.797826
Jitter:DDP	-0.0762127	0.0971499	-0.100488	0.990276	0.922913	1	0.957319	1	0.759555
MDVP:Shimmer	-0.0983737	0.00228123	-0.144543	0.769063	0.703322	0.759581	0.797826	0.759555	1
MDVP:Shimmer(dB)	-0.0737425	0.0434652	-0.119089	0.804289	0.716601	0.790652	0.839239	0.790621	0.987258
Shimmer:APQ3	-0.0947171	-0.00374325	-0.150747	0.746625	0.697153	0.744912	0.76358	0.744894	0.987625
Shimmer:APQ5	-0.0706818	-0.00999678	-0.101095	0.725561	0.648961	0.709927	0.78678	0.709907	0.982835
MDVP:APQ	-0.0777738	0.00493698	-0.107293	0.758255	0.648793	0.737455	0.804139	0.737439	0.950083
Shimmer:DDA	-0.0947316	-0.00373289	-0.150737	0.746635	0.69717	0.744919	0.763592	0.744901	0.987626
NHR	-0.0219808	0.163766	-0.10867	0.906959	0.834972	0.919521	0.844604	0.919548	0.722194
HNR	0.0591444	-0.0248931	0.210851	-0.728165	-0.65681	-0.721543	-0.73151	-0.721494	-0.835271
status	-0.383535	-0.166136	-0.3802	0.27822	0.338653	0.266668	0.288698	0.266646	0.36743
RPDE	-0.383894	-0.112404	-0.400143	0.360673	0.441839	0.34214	0.333274	0.342079	0.447424
DFA	-0.446013	-0.343097	-0.0504063	0.0985724	0.175036	0.064083	0.196301	0.0640264	0.159954

Why We need it?

Parkinson's Disease is a chronic and progressive neurodegenerative disorder that affects millions of people worldwide. Early detection of Parkinson's Disease is crucial for effective treatment and management of the disease, as early intervention can help slow the progression of symptoms and improve quality of life for patients. However, current methods for Parkinson's Disease diagnosis are often subjective and rely on clinical observation, which can lead to misdiagnosis and delayed treatment.

Machine learning has the potential to significantly improve the accuracy and efficiency of Parkinson's Disease diagnosis. By training a machine learning model on a dataset of voice recordings from individuals with and without Parkinson's Disease, the model can learn to recognize patterns in the data that are indicative of the disease. This can lead to more objective and accurate diagnosis of Parkinson's Disease, which can ultimately improve patient outcomes. Overall, the development of a machine learning model for the detection of Parkinson's Disease is needed to improve the accuracy and efficiency of diagnosis, and to provide patients with earlier intervention and better treatment options.

Project objective

Our 4 Major Objectives of our project are:-

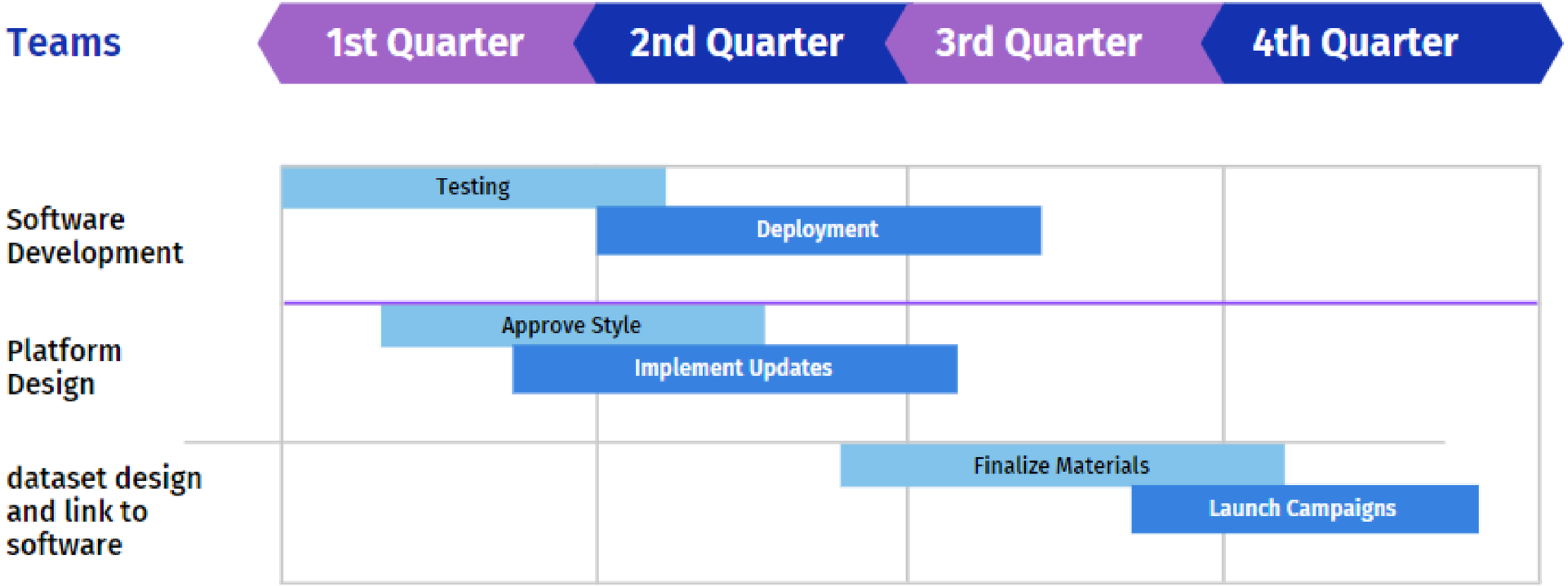
- To develop a machine learning model for the detection of Parkinson's Disease based on voice features.
- To train the machine learning model on a dataset of voice recordings from individuals with and without Parkinson's Disease.
- To evaluate the performance of the machine learning model using metrics such as accuracy, precision, recall, and F1 score.
- To compare the performance of the machine learning model to existing methods for Parkinson's Disease diagnosis.
- To identify the most important voice features for Parkinson's Disease detection using feature selection techniques.
- To develop a user-friendly application for the deployment of the machine learning model.

Facilities required for proposed work

The facilities required for the development of this project include both software and hardware.
The software required for this project includes Python programming language, machine learning libraries such as scikit-learn, TensorFlow, and Keras, and database management systems such as MySQL. Additionally, software for developing the user-friendly application, such as HTML, CSS, and JavaScript, will be required.
The hardware required for this project includes a computer with sufficient processing power and memory to handle the data preprocessing and machine learning algorithms. A microphone or other audio recording device will also be required to collect voice recordings for training and testing the machine learning model.
Overall, the facilities required for this project are readily available and accessible, and can be easily obtained and installed. The software and hardware requirements are standard for machine learning and software development projects, and can be easily met by most modern computers and devices.

Timeline

TIMELINE



LITERATURE REVIEW

S.No	PAPER	YEAR	AUTHOR	APPROACH	PERFORMANCE ANALYSIS
1	"Detection of Persons with Parkinson's Disease by Acoustic, Vocal, and Prosodic Analysis"	2011 IEEE	"Tobias Bocklet , Elmar N¨oth , Georg Stemmer , Hana Ruzickova , Jan Rusz"	SVM - algorithm Prosodic -base feature	recognition rate 90.5% AUC 0.97
2	"Using Machine Learning to Diagnose Parkinson's Disease from Voice Recordings"	2017 IEEE	" Akshaya Dinesh, Jennifer He"	gradient boosted regression trees	91-95% accuracy
3	"Parkinson's Disease Diagnosis Using Machine Learning and Voice"	November 2018	"Timothy J. Wroge, Yasin O¨ zkanca, Cenk Demiroglu, Dong Si, David C. Atkins, Reza Hosseini Ghomi"	Gradient boosted classifier.	85% accuracy
4	"Deep Learning Approach to Parkinson's Disease Detection Using Voice Recordings and Convolutional Neural Network Dedicated to Image Classification"	2019 IEEE	"Marek Wodziński, Andrzej Skalski, Daria Hemmerling, Juan Rafael Orozco-Arroyave, Elmar N¨oth"	ResNet algorithm	90% accuracy
5	"Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease	August 9, 2017	"Betul Erdogdu Sakar, Gorkem Serbes, C. Okan Sakar"	SVM	95% accuracy
6	"Suitability of dysphonia measurements for telemonitoring of Parkinson's disease"	11 Sep 2008	"Max A. Little1, Member IEEE, Patrick E. McSharry, Senior Member	SVM	91.4% accuracy

			Eric J. Hunter, Jennifer Spielman, Lorraine O. Ramig."		
7	"Voice-Based Detection of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study"	23 August 2019	"Iqra Nissar1, Danish Raza Rizvi1, Sarfaraz Masood, Aqib Nazir Mir"	XGBoost	95% accuracy
8	"Prediction of Parkinson's disease using Machine Learning Techniques"	June 2018	"Kirti Sharma"	Random Forest	95% accuracy
9	"Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis"	2019	"Amit Kumar Patra Ratula Ray , Azian Azamimi Abdullah , Satya Ranjan Dash"	Ensemble Modelling	81-85% accuracy
10	"High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing"	November 2017	"Warwick R. Adams"	Ensemble modeling	sensitivity -92 to 100%, a specificity- 95 to 100%, AUC -f between 0.97 and1.00.

THANK YOU