RESEARCH ARTICLE OPEN ACCESS

Parkinson Disease Prediction Using Machine Learning Algorithm Models

Dr K. Sailaja MCA, M.Tech , M.Phil, Ph.D [1], M.jayaprakash Reddy [2]

[1] Professor & HOD, Department of Computer Applications

[2] Student, Department of Computer Applications

[1], [2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT

Parkinson's disease (PD) is the second commonest late life neurodegenerative disease after Alzheimer's disease. It is prevalent throughout the world and predominantly affects patients above 60 years old. Diagnosis of Parkinson's disease (PD) is commonly based on medical observations and assessment of clinical signs, including the characterization of a variety of motor symptoms. However, traditional diagnostic approaches may suffer from subjectivity as they rely on the evaluation of movements that are sometimes subtle to human eyes and therefore difficult to classify, leading to possible misclassification. In the meantime, early non-motor symptoms of PD may be mild and can be caused by many other conditions. Therefore, these symptoms are often overlooked, making diagnosis of PD at an early stage challenging. An innovative deep-learning technique is introduced to early uncover whether an individual is affected with PD or not based on premotor features. Specifically, to uncover PD at an early stage, several indicators have been considered in this study. Deep learning methods based on relatively small data including 183 healthy individuals and 401 early PD patients shows the superior detection performance of the designed XGBoost model, which achieves the highest accuracy, 80-90% on average. Besides detecting the PD, we also do comparison between the SVM and XGBoost model on basis of accuracy.

Keywords: - XGBoost Model, Accuracy, Parkinson, Diagnosis, Alzheimer disease.

I. INTRODUCTION

Parkinson disease mainly effects central nervous system and is observed to be affected on many people globally. Most of the people suffering with PD are observed to be physically and emotionally draining. They even feel depressed, trouble concentrating on things, painful spasms etc. PD has a large spectrum of clinical features ranging from motor to nonmotor symptoms. Some of the motor symptoms are hypophonic speech, rigidity, resting tremor. Non-motor symptoms are as hallucinations, depression, constipation, sleeping disorders, cognitive impairment, and impulse control disorders. Non motor symptoms show sickness than motor symptoms [1,3]. Most of the cases, physicians find it difficult to envisage whether a given patient is already affected by the disease or is expected to develop the Parkinson's disease[7]. To conquer this, development of some computing model must be done that evaluates and summarizes the data of a given patient and predicts with adequate accuracy where he/she will have development of PD. Most of the PD patients are observed with symptoms called voice impairment which is known as dysphonia. There are several measures related to dysphonia, out of which voice related problem can be used to assess the patients at various stages[14]. This paper is a survey of prediction of PD using Machine learning and Deep learning techniques that generated good models and potency of those algorithms in terms of accuracies achieved, also about different methodologies applied.

Parkinson's disease symptoms can be different for everyone. Early signs are mild that goes unnoticed. Symptoms usually begin on one side of your body and gets worsen on that side, afterwards it affects both the sides. Parkinson's symptoms may include Tremor, Slowed movement, Rigid muscles, Impaired posture and balance, Loss of automatic movements, Speech changes, Writing changes. The Parkinson's disease is due to a loss of neurons that produce a chemical messenger in the brain called dopamine. when there is a decrease in level of the amino acid named dopamine it leads to the abnormal brain activity, which leads to Parkinson's disease. The cause of Parkinson's disease is still a question mark, but several factors appear to play a role, including: Genes, Environmental, Triggers. As a result people suffer from this disease for many years before

diagnosis. The estimated results have shown that there are 7-10 million people are affected by Parkinson's disease worldwide. People with age above 50 are the one's who has the higher possibility of getting Parkinson's disease but still an estimated 4 percentage of people who are under the age 50 are diagnosed with Parkinson's disease. There is no cure or prevention for PD. However, the disease can be controlled in early stage. The data mining techniques is used as a effective way for early detection and diagnosis of the disease.

Data mining techniques in medicine is a research area that combines sophisticated representational and computing techniques with the insights of expert physicians to produce tools for improving healthcare.Data mining is a statistical method for finding hidden patterns in datasets by constructing predictive or classification models that can be learned from past experience and applied in future cases, so there is a need for a more accurate, objective means of early detection, ideally one which can be used by individuals in their home setting.

II. RELATEDWORKS

Marwa Almasoud, Tomas E Ward Proposed Parkinson Disease Gait Classification Based On Machine Learning Approach in 2013. This study discussed the ability of two machine classifiers namely Artificial Neural Network (ANN) and Support Vector Machine (SVM) in distinguishing gait pattern during self-selected speed walking due to the effect of motor Parkinson Disease (PD). There are three gait parameters that is utilized as features in classifying PD gait and normal subjects namely basic spatiotemporal, kinematic and kinetic. Firstly, the input features are pre-processed using two types of normalization technique specifically intra group as well as inter group normalization. Additionally, all the three features are classified solely followed by implementation of data fusion. Then, the effectiveness of the features vectors to identify PD patients or vice versa is evaluated based as inputs to both classifiers. Initial findings showed that basic spatiotemporal solely as feature vectors based on intra group normalization technique contributed perfect classification for both ANN and SVM as classifiers.

Sanghee Moon, Hyun-Je Song, Vibhash D. Sharma, Kelly E. Lyons, Rajesh Pahwa proposed Classification Of Parkinson's Disease And Essential Tremor Based On Balance And Gait Characteristics From Wearable Motion Sensors Via Machine Learning in the year 2020. Parkinson's disease (PD) and essential tremor (ET) are movement disorders that can have similar clinical characteristics including tremor and gait difficulty. These disorders can be misdiagnosed leading to delay in appropriate treatment. The aim of the study was to determine whether balance and gait variables obtained with wearable inertial motion sensors can be utilized to differentiate between PD and ET using machine learning. Additionally, we compared classification performances of several machine learning models.
This retrospective study included balance and gait variables collected during the instrumented stand and walk test from people with PD (n = 524) and with ET (n = 43). Performance of several machine learning techniques including neural networks, support vector machine, k-nearest neighbor, decision tree, random forest, and gradient boosting, were compared with a dummy model or logistic regression using F1-scores. Machine learning models classified PD and ET based on balance and gait characteristics better than the dummy model (F1-score = 0.48) or logistic regression (F1-score = 0.53). The highest F1-score was 0.61 of neural network, followed by 0.59 of gradient boosting, 0.56 of random forest, 0.55 of support vector machine, 0.53 of decision tree, and 0.49 of k-nearest neighbor. This study demonstrated the utility of machine learning models to classify different movement disorders based on balance and gait characteristics collected from wearable sensors. Future studies using a well-balanced data set are needed to confirm the potential clinical utility of machine learning models to discern between PD and ET.

Basetty Mallikarjuna, R. Viswanathan and Bharat Bhushan Naib proposed Feedback-Based Gait Identification Using Deep Neural Network classification in the year 2020. Identification of gait plays a major role in the healthcare industry, recognition of a gait having different angles, identification of abnormalities is a challenging task, to detect the abnormal person identification contains improper pattern style, human limbs, walking pattern, etc... A normal person has a correct pattern, an

abnormal person has an irregular pattern. This paper provides the identification of the lean angle and ramp angle [19] of irregular patterns on three abnormalities such as Parkinson gait, Hemiplegic gait, and Neuropathic gait [18] by using deep neural network (DNN) without clinical observation by using DNN classification with feedback-based verification of trained features with query features of abnormal identification of trained features with query features. This paper concludes the gait abnormalities based on lean angle and ramp angle.

Rana Zia Ur Rehman, Silvia Del Din, Yu Guan, Alison J. Yarnall proposed Selecting Clinically Relevant Gait Characteristics For Classification Of Early Parkinson's Disease in the year 2020. Parkinson's disease (PD) is the second most common neurodegenerative disease; gait impairments are typical and are associated with increased fall risk and poor quality of life. Gait is potentially a useful biomarker to help discriminate PD at an early stage, however the optimal characteristics and combination are unclear. In this study, we used machine learning (ML) techniques to determine the optimal combination of gait characteristics to discriminate people with PD and healthy controls (HC). 303 participants (119 PD, 184 HC) walked continuously around a circuit for 2-minutes at a self-paced walk. Gait was quantified using an instrumented mat (GAITRite) from which 16 gait characteristics were derived and assessed. Gait characteristics were selected using different ML approaches to determine the optimal method (random forest with information gain and recursive features elimination (RFE) technique with support vector machine (SVM) and logistic regression). Five clinical gait characteristics were identified with RFE-SVM (mean step velocity, mean step length, step length variability, mean step width, and step width variability) that accurately classified PD. Model accuracy for classification of early PD ranged between 73–97% with 63–100% sensitivity and 79–94% specificity. In conclusion, we identified a subset of gait characteristics for accurate early classification of PD. These findings pave the way for a better understanding of the utility of ML techniques to support informed clinical decision-making.

Milla Juutinen,Justin Zhu, Cassia Wang proposed Perkinson's Disease Detection From 20-Step Walking Tests Using Inertial Sensors Of Smartphone in the year 2020. Parkinson's disease (PD) is a neurodegenerative disease inducing dystrophy of the motor system. Automatic movement analysis systems have potential in improving patient care by enabling personalized and more accurate adjust of treatment. These systems utilize machine learning to classify the movement properties based on the features derived from the signals. Smartphones can provide an inexpensive measurement platform with their built-in sensors for movement assessment. This study compared three feature selection and nine classification methods for identifying PD patients from control subjects based on accelerometer and gyroscope signals measured with a smartphone during a 20-step walking test. Minimum Redundancy Maximum Relevance (mRMR) and sequential feature selection with both forward (SFS) and backward (SBS) propagation directions were used in this study. The number of selected features was narrowed down from 201 to 4–15 features by applying SFS and mRMR methods. From the methods compared in this study, the highest accuracy for individual steps was achieved with SFS (7 features) and Naive Bayes classifier (accuracy 75.3%), and the second highest accuracy with SFS (4 features) and k Nearest neighbours (accuracy 75.1%). Leave-one-subject-out cross-validation was used in the analysis. For the overall classification of each subject, which was based on the majority vote of the classified steps, k Nearest Neighbors provided the most accurate result with an accuracy of 84.5% and an error rate of 15.5%. This study shows the differences in feature selection methods and classifiers and provides generalizations for optimizing methodologies for smartphone-based monitoring of PD patients. The results are promising for further developing the analysis system for longer measurements carried out in free-living conditions.

III. PROPOSED SYSTEM ARCHITECTURE

Parkinson's disease (a neurodegenerative disorder) that causes the patients' motor abilities to degrade over time due to the damage caused to the dopamine-generating brain cells. Shaking, trouble moving, behavioral disorders, dementia, and

depression are some of the results of this disorder. The primary motor conditions are referred to as ”Parkinsonism,” or a ”Patient with Parkinson’s Disease.” One of the most common symptom that can be recognized by studying the patients’ voice data is changes in their voice. The patient’s speech stutters and becomes increasingly impacted as the disease progresses. a single unit for classification and feature selection models. Deep learning is being used in this paper to classify the patient’s voice data into ”extreme” and ”not severe” categories. The two UPDRS (Unified Parkinson’s Disease Rating Scale) scores - total UPDRS and motor UPDRS - were used as assessment criteria in this study. The motor UPDRS assesses the patient’s motor capacity in the scale of 0-108, while the total UPDRS assesses the patient’s overall ability and its score range from 0-176. Deep learning has risen in importance as a method for analysing unstructured data such as speech and audio signals. Multiple layers of neurons are often used in deep neural networks, these layers are stacked as a single unit for classification and feature selection models. Prognosis and progression of Parkinson’s disease is a critical question among the clinicians since there is a disparity of parameters taken into the diagnostic consideration thereby making the decision process difficult. Different datasets have been independently explored and applied through deep learning to analyze the incidence of occurrence and progression of the disease. The present paper is an updated report of the types of deep Learning algorithms which have gained prominence within a span of last 5 years (2015-2019). Further it highlights the use of hybrid intelligence models to improve the prediction accuracy and sensitivity over standalone SVM and XGBoost methods which is shown in Fig.1.

By using Machine learning techniques, the problem can be solved with minimal error rate. The voice dataset of Parkinson's disease from the UCI Machine learning library is used as input. Also our proposed system provides accurate results by integrating spiral drawing inputs of normal and Parkinson’s affected patients. We propose a hybrid and accurate results analyzing patient both voice and spiral drawing data’s. Thus combining both the results, the doctor can conclude normality or abnormality and prescribe the medicine based on the affected stage.

PROPOSED SYSTEM ADVANTAGES

- Complexity is less compared to previous process
- Ability to learn and extract complex features.
- Accuracy is good
- With its simplicity and fast processing time, the proposed algorithm gives better execution time.

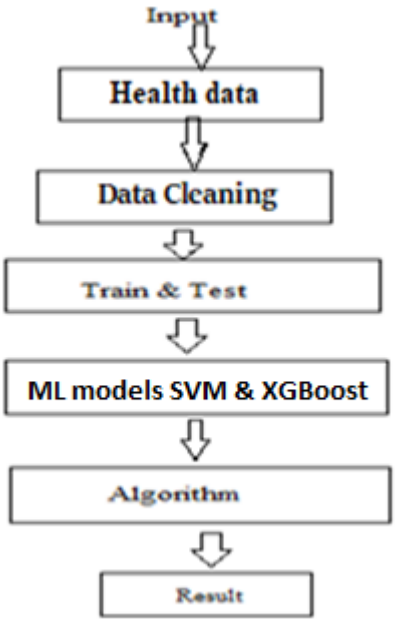

Fig.1 Proposed system Architecture

IV. RESULTS AND DISCUSSION

The output screens obtained after running and executing the system are shown from Fig.2 to Fig.10.

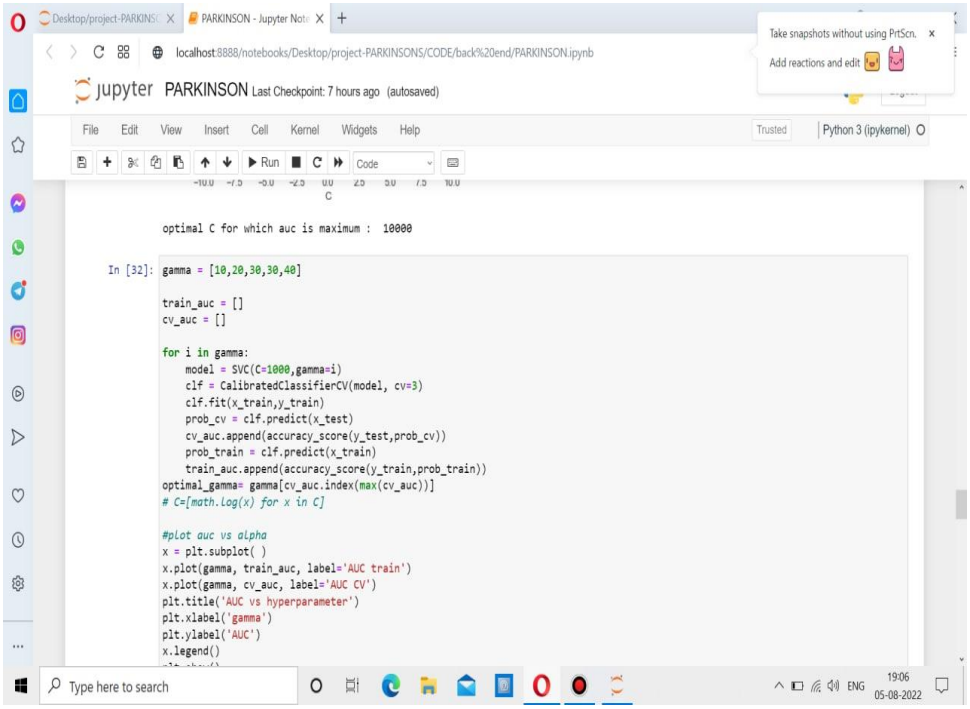

Fig.2 Plot AUC versus Alpha

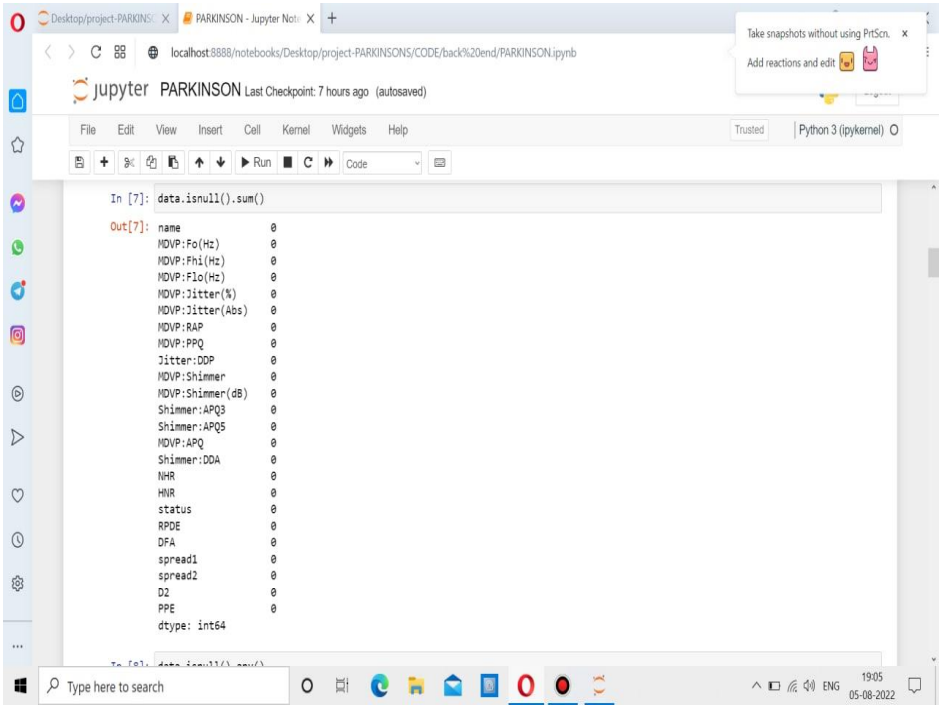

Fig. 3 data sum

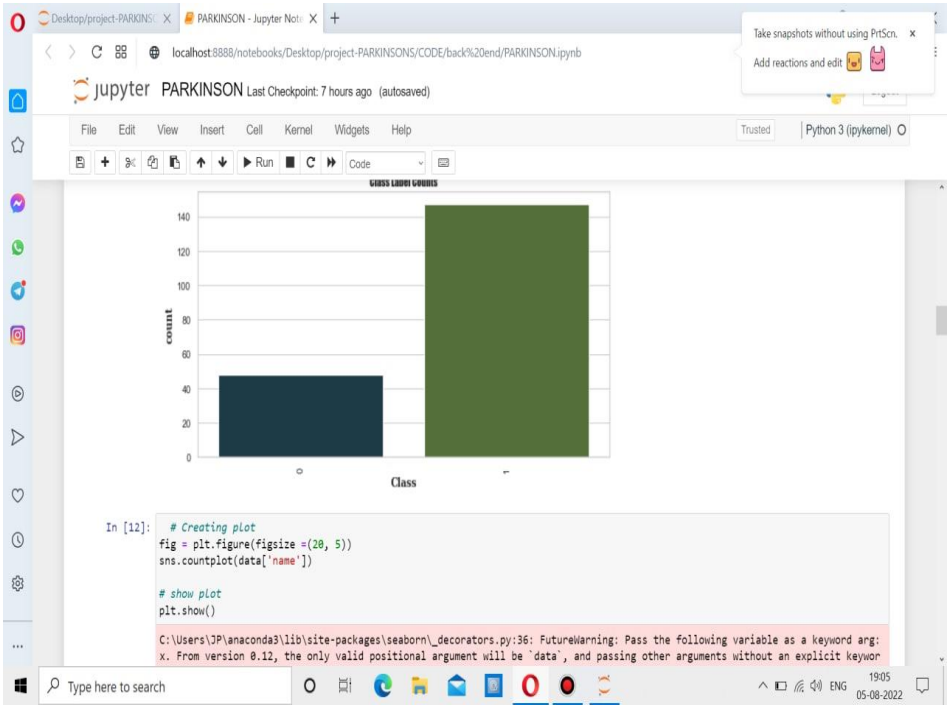

Fig.4 Creating and showing plot

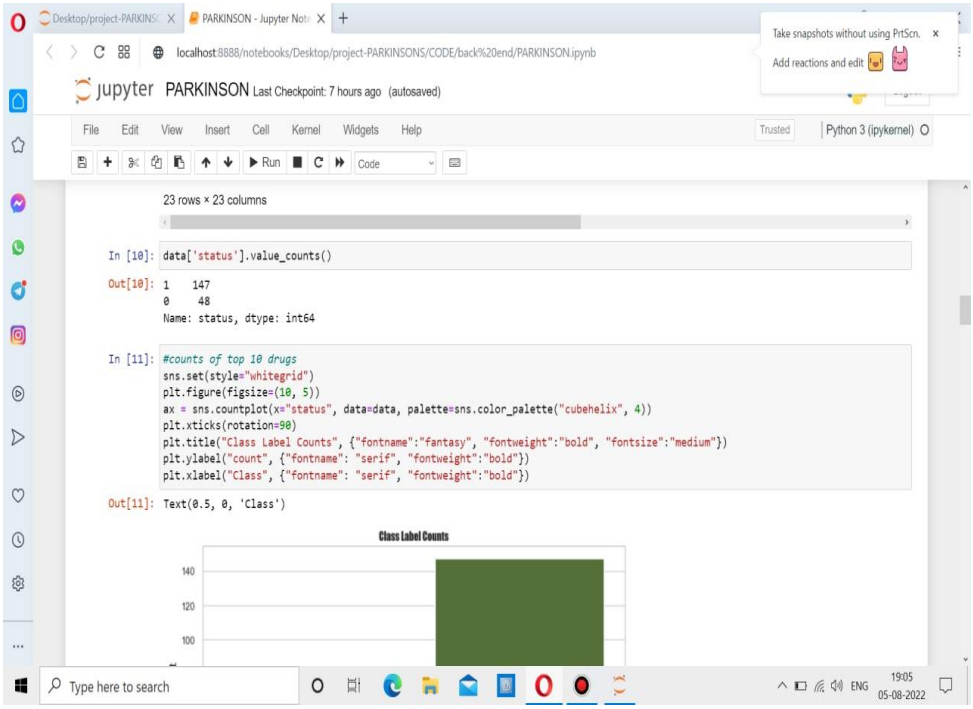

Fig.5 Class label counts

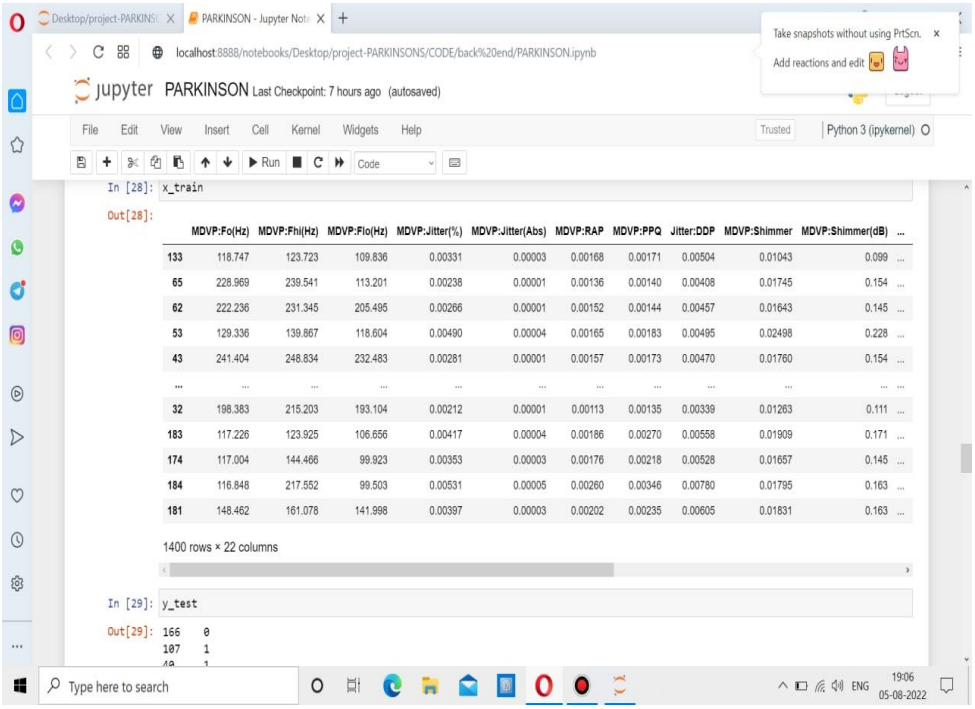

Fig.6 Training model

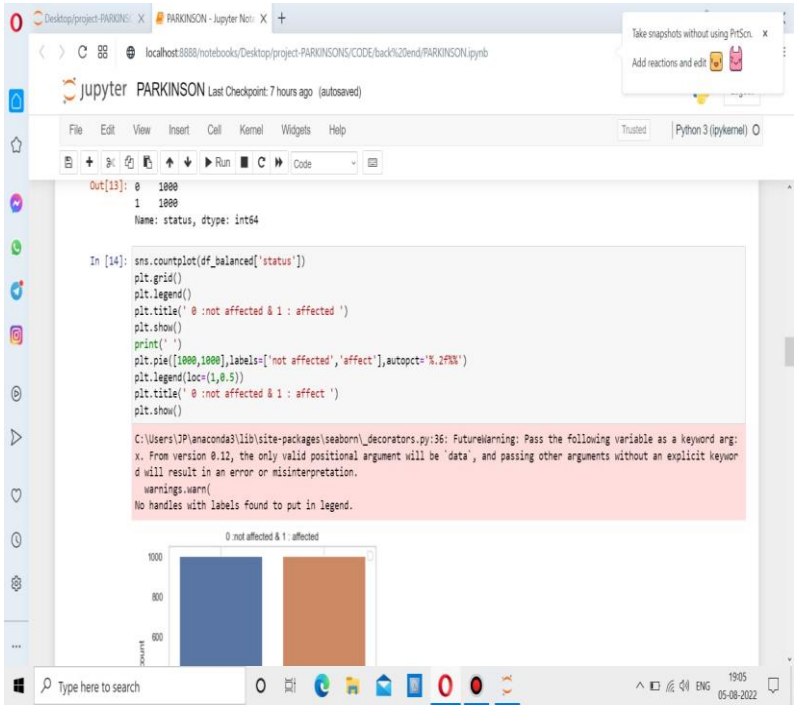

Fig.7 Count Plot

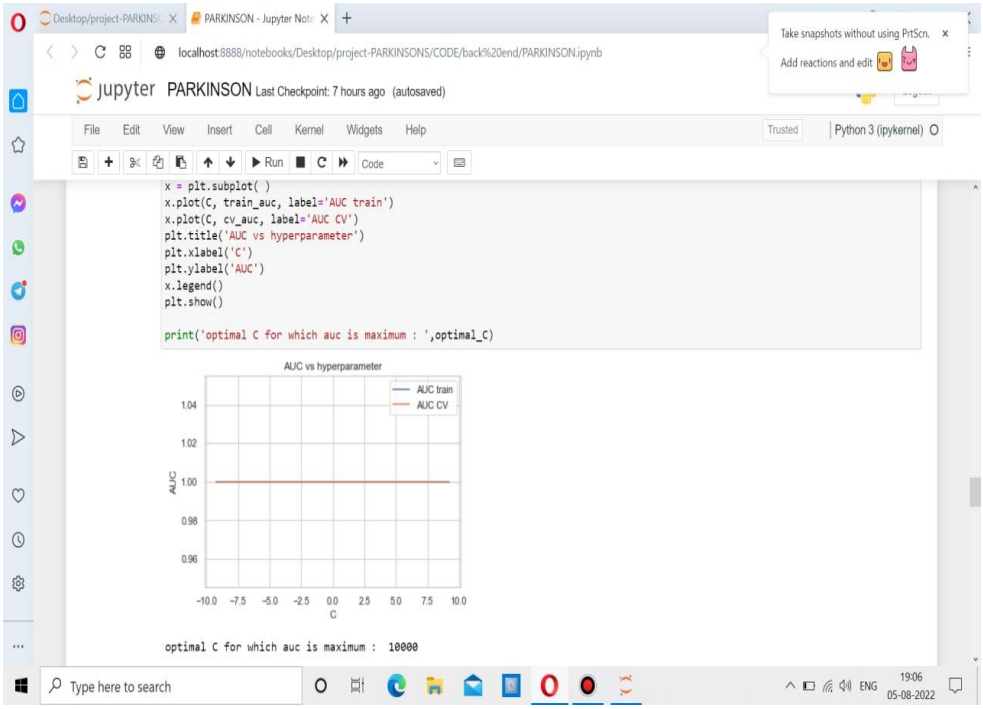

Fig.8 optimal 'C' value

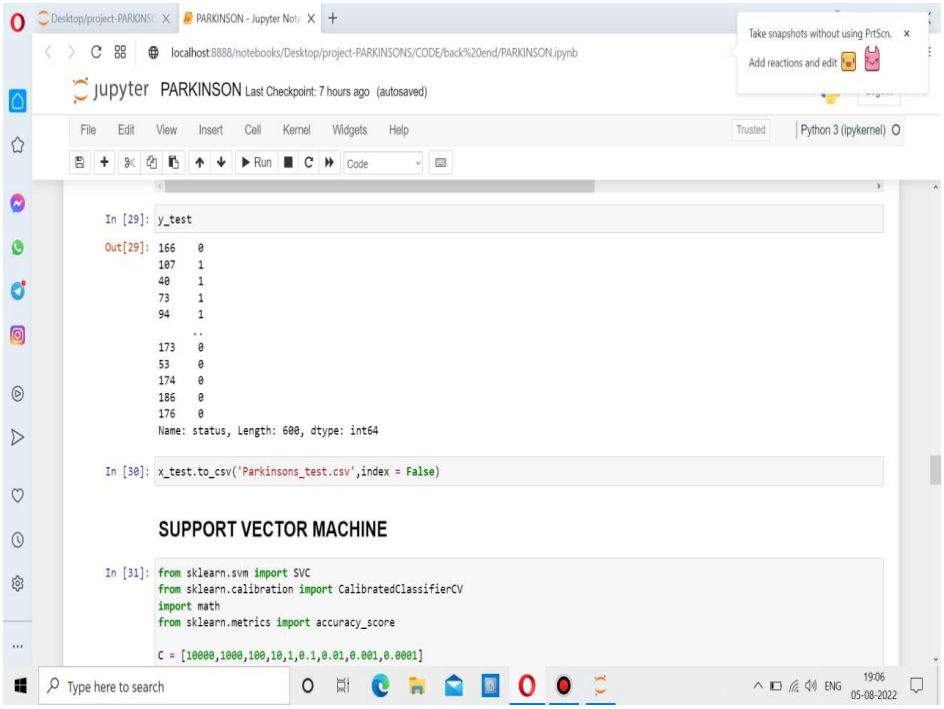

Fig.9 Testing Model

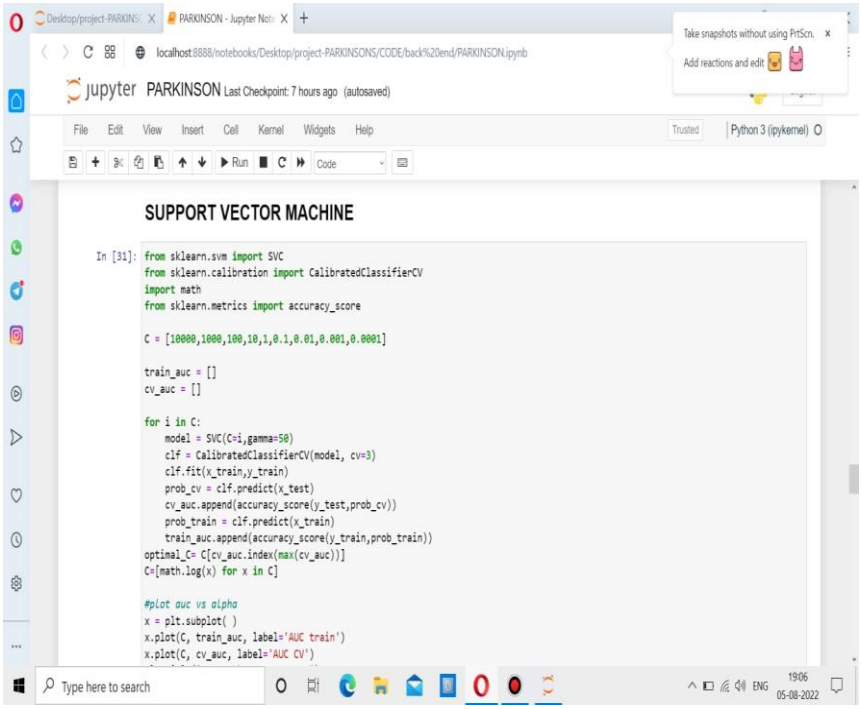

Fig.10 Support vector Machine

V. FUTURE SCOPE AND CONCLUSION

This paper is an effort to present broad review about Parkinson disease diagnosis system that have applied various deep learning techniques. This project aimed to cover a broader space of imaging and machine learning technologies for mental illness diagnostics such that researchers in the field could readily identify the state of the art in the domain. Moreover, we emphasize the importance of early detection and prediction of Parkinson's disease, such that treatment and support can be provided to patients as soon as possible. It can be identified that maximum of all ML techniques used by various authors worked better but developing a very faster classifier using novel architecture of ML combined with specific approach may work better. To achieve this, we try to implement convolution neural network with different number of ML and number of nodes in future and compare all the accuracies. In future work, we can focus on different techniques to predict the Parkinson disease using different datasets. In this research, we using binary attribute (1- diseased patients, 0-non-diseased patients) for patient's classification. In the future we will use different types of attributes for the classification of patients and also identify the different st ages of Parkinson's disease.

REFERENCES

[1] R. Prashanth and S. D. Roy, ''Early detection of Parkinson's disease through patient questionnaire and predictive modelling,'' Int. J. Med. Informat., vol. 119, pp. 75–87, Nov. 2018.

[2] N. Singh, V. Pillay, and Y. E. Choonara, ''Advances in the treatment of Parkinson's disease,'' Prog. Neurobiol., vol. 81, pp. 29–44, Jan. 2007.

[3] H. Gunduz, ''Deep learning-based Parkinson's disease classification using vocal feature sets,'' IEEE Access, vol. 7, pp. 115540–115551, 2019.

[4] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, ''Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests,'' IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 884–893, Apr. 2010.

[5] S. Lahmiri and A. Shmuel, ''Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine,'' Biomed. Signal Process. Control, vol. 49, pp. 427–433, Mar. 2019.

[6] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, ‘‘Automatic detection of Parkinson’s disease based on acoustic analysis of speech,’’ Eng. Appl. Artif. Intell., vol. 77, pp. 148–158, Jan. 2019.

[7] J. Kumar, and A. K. Singh, "Dynamic resource scaling in cloud using neural network and black hole algorithm," 2016 Fifth International Conference on Eco-friendly Computing and Communication Systems (ICECCS), Bhopal, 2016, pp. 63-67.

[8] J. Kumar, A.K. Singh, “Workload prediction in cloud using artificial neural network and adaptive differential evolution,” Future Generation Computer Systems, 2018 Apr 1;81:41-52.

[9] A. Singh, and J. Kumar, “Secure and Energy Aware Load Balancing Framework for Cloud Datacenter Networks,” Electronics Letters, 2019 Mar 8, in press.

[10] H. Taneja, and A.K. Singh, “Preserving Privacy of Patients Based on Re-identification,” Risk. Procedia Computer Science, 2015 Jan 1;70:448-54.