

AN INTEGRATIVE APPROACH TO DIAGNOSING PARKINSON'S DISEASE USING ENSEMBLED MACHINE LEARNING MODELS

Abstract

In the early stages of Parkinson's disease (PD), a neurodegenerative disorder that affects the brain's neurological, physiological, and behavioural systems, a correct diagnosis can be challenging due to the presence of subtle variances. Slow movement, or bradykinesia, is a common sign of this condition. The onset of symptoms occurs around midlife, and the condition progresses steadily worse with age. Changes in speech are one of the early indicators of Parkinson's disease. This study proposed the use of supervised classification algorithms like support vector machine (SVM), naive Bayes (NB), k-nearest neighbour (K-NN), and artificial neural network (ANN) for diagnosing subjective diseases.

To enhance resilience beyond the capabilities of a single model, a method known as ensembling has been implemented. Ensembling involves training numerous models instead of relying on a single model, then subsequently combining the predictions generated by these models. This research evaluates four distinct ensemble algorithms, including Adaboost, bagging, majority voting and weighted voting. To assess the efficacy of the proposed methodology, a dataset comprising speech signal collection includes 195 biomedical sounds, of which there are 147 phonetics for persons with Parkinson's disease (PD) and 48 for healthy people. Through analysis, it was determined that the implementation of an

Authors

Dr. A Deepa

Associate Professor
Sathyabama Institute of Science and
Technology
Chennai, India.
adeepa.cse@sathyabama.ac.in

Dr. S. Bangaru Kamatchi

Assistant professor
Sathyabama Institute of Science and
Technology
Chennai, India.
bangarukamatchi.cse@sathyabama.ac.in

Dr. M. P. Vaishnave

Assistant professor
Sathyabama Institute of Science and
Technology
Chennai, India.
vaishnave.cse@sathyabama.ac.in

ensemble approach utilising the majority voting technique, in conjunction with MLP, SVM and decision tree classifiers, yielded superior results. Specifically, the ensemble approach achieved an accuracy rate of 95.7%, surpassing the performance of the individual classifiers.

Keywords: Parkinson's disease; machine learning, disease prediction; neurodegenerative disorder; adaboost; bagging; majority voting; soft voting; ensemble;

I. INTRODUCTION

Parkinson's disease, also known as tremor, is caused by a decrease in the amounts of dopamine that are present in the brain. This decrease causes harm to a person's motor functions, as well as their physical functioning. It is one of the most widespread illnesses in the entire planet. These lesions cause neurological signs and symptoms to appear at irregular intervals, which get more severe as the disease advances [1]. This disorder is more prevalent among the older population because ageing produces changes in our brains, including the loss of synaptic connections as well as alterations in neurotransmitters and neurohormones. The neurons in an individual's body start to degenerate and lose their ability to be replicated as time goes on. Because the repercussions of neurological disorders and the declining levels of dopamine in the patient's body show up gradually, it is difficult to notice them until the patient's condition is severe enough to necessitate medical care[2]. However, various people experience varying degrees of symptoms and severity levels of the condition. This condition is characterised by major symptoms such as difficulty speaking, loss of memory for recent events, loss of balance, and posture that is imbalanced [3].

According to a research published by the World Health Organisation (WHO) in 2019 [4], this condition affects 8.5 million people on an annual basis and may be found all over the world. The likelihood of getting this condition increases with the patient's age; nonetheless, there are now 4% of patients under the age of 50 in the world. This condition is the second most prevalent form of neurodegeneration in the world, behind Alzheimer's disease, and it affects millions of individuals [5]. Because there is currently no cure for this illness, medical professionals can only provide patients assistance in managing the symptoms of the condition [6]. However, a definitive diagnostic test for this disease does not exist, and instead, the diagnosis relies heavily on the patient's previous health conditions [7]. Because invasive procedures are frequently utilised for diagnosis and therapy, which are not only costly but also demanding [8], it appears that a method of diagnosing this condition that is relatively basic and accurate is highly significant.

Researchers have spent the better part of the last several decades looking into machine learning (ML), a subfield of artificial intelligence (AI), as a potential new method for diagnosing this condition. When clinical staff uses ML in conjunction with more traditional diagnostic indicators, they may be better able to identify individuals as having this condition [9]. Because walking is the activity that people engage in the most often in their day-to-day lives, it has been connected to a variety of neurological and physical conditions. Data on gait (also known as mobility) has, for instance, allowed for the identification of this disease. Approaches to gait analysis have a number of benefits, including the fact that they are not invasive and have the potential to be utilised to a large extent in residential settings [10, 11]. Very few academics have attempted to integrate several machine learning algorithms to create a process that is fully automated and can be carried out offline [12, 13]. In addition, those with the condition who are in the early stages of the disease may have difficulties communicating. These conditions include dysphonia (a lack of vocal fluency), repeated echoes (a small selection of aural changes), and hypophonia (a discord in the vocal muscles). It is possible for a computing unit to identify and assess information that is derived from human auditory emissions [14].

The study is structured as follows: Section 2 provides an overview of the survey conducted in previous research studies. Section 3 provides an in-depth analysis of the methods employed in order to accomplish the stated aim. Section 4 of this study focuses on the analysis of the experiment and presents the obtained data. Section 6 provide the conclusion of research and represents the further level of enhancement in research.

II. LITERATURE SURVEY

This paper research is distinguished from other studies by the development of diagnostic systems with a variety of approaches and tools that are capable of successfully analysing audio data and distinguishing between persons with Parkinson's disease and healthy people with a high degree of precision. In their study, M. AI-Sarem et al. [15] proposed a methodology for the identification of Parkinson's disease utilising several techniques such as random forest, extreme gradient boosting (XGBoost), and CatBoost, with the aim of enhancing the overall accuracy of the detection process. The researchers analyse the key attributes of each ensemble classification method, namely random forest, extreme gradient boosting (XGBoost), and categorical boosting (CatBoost).

In their study, T. J. Wronge et al. [16] introduced a Voice Activation Detection (VAD) technique aimed at predicting Parkinson's disease. The process involved the extraction of raw audio data, followed by the elimination of background noise. Subsequently, the processed audio was subjected to two separate feature-extraction algorithms. Ultimately, an algorithm is employed for the purpose of machine learning. The study conducted by K.R. Wan et al. [17] involved the analysis of research articles that might be employed in the field of machine learning (ML) for the purpose of selecting functions (FS) in brain surgery. A machine learning (ML) based methodology is employed to identify the precise anatomical location that requires intervention during surgical procedures targeting Parkinson's disease. The primary emphasis of this study pertains to studies that are carried out subsequent to the diagnosis of Parkinson's disease.

In their study, Cavallo et al. [18] made an effort to predict the occurrence of Parkinson's disease by utilising motion data obtained from the upper limbs of individuals. The experimental participants, consisting of both individuals with Parkinson's disease and healthy individuals, had a device surgically implanted into their upper limbs. They were then given instructions to carry out a variety of exercises. In order to get parameters, a comprehensive analysis of spatiotemporal and frequency data was conducted, followed by the implementation of several supervised learning methodologies for the purpose of categorization. In their study, J.S. Almeida et al. [19] utilised a range of techniques for feature extraction and machine learning in order to identify Parkinson's disease (PD). It has been determined that phonation represents the most efficacious activity for the detection of Parkinson's disease. The study investigated the classification performance of K-NN, Multilayer Perceptron (MLP), Optimum Path Forest, and Support Vector Machines (SVM) in the research analysis. In a study conducted by Parisi et al. [20], artificial neural networks were employed to minimise speech features in order to facilitate the detection of Parkinson's disease using machine learning techniques. Support Vector Machines (SVM) were employed for the purpose of data classification. In addition to utilising ML algorithms to analyse motion or voice information, researchers have also discovered Parkinson's disease (PD) through the analysis of handwriting exercises [21].

Although, ML-based PD diagnosis has been reported to achieve high classification rates in the existing research. The authors either used a substantial quantity of features, resulting in longer computation time, or encountered challenges in extracting the features, even when employing only a limited number of them. Consequently, the duration of computation is indirectly prolonged. The objective of this research is to save computational time by employing a reduced set of useful features through a streamlined feature extraction procedure. The characteristics are obtained from speech signals, which simplify the process of collecting these features compared to other methodologies found in the literature, such as MRI-based or motion-based methods. The primary contributions of this study are centred around the prediction of early Parkinson's disease using machine learning techniques.

III. PROPOSED METHODOLOGY

In this part, the methodologies for feature extraction, the traditional classifiers used for assembling, which are also referred to as base classifiers, and the suggested ensemble classifier are discussed. In the first step of the process, the features from the dataset are extracted with the help of the Speed up Robust Features methods. After that, the values of the extracted features are input into the base classifiers, which can include things like Support Vector Machine, Decision Tree, Logistic Regression, Random Forest, Multilayer Perceptron, Naive Bayes, and k-Nearest Neighbours. In order to get a higher level of accuracy in PD disease diagnosis, the three best base classifiers are combined using the ensemble method.

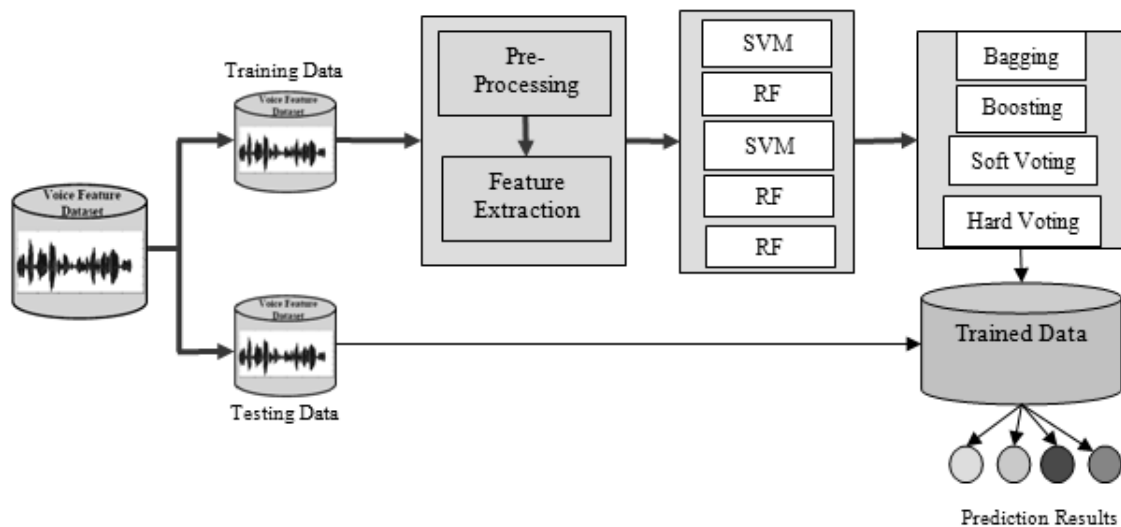


Figure 1: Architecture of Integrated Approaches of Parkinson Disease Diagnosis

The proposed PD diagnosis approach is consisting of following stages:

- Dataset Collection stage
- Pre-processing stage
- Feature extraction stage
- Traditional Classification
- Ensembled Classification (Improving the Accuracy).

1. Dataset Collection Stage

The Parkinson's disease (PD) dataset that was utilised for the early identification of PD in this article is based on speech sounds. It was generated and provided to the UCI Machine Learning Repository by Max Little of Oxford University [44]. Numerous medical professionals agree that the data set is among the most effective ones ever gathered, organised, and analysed. Quite a few researchers have created automated methods, and they have tested them on this dataset. It is still the goal of a significant number of researchers and other individuals interested in the diagnosis of PD at an earlier stage. The speech signal collection includes 195 biomedical sounds, of which there are 147 phonetics for persons with Parkinson's disease (PD) and 48 for healthy people [22]. The first table displays 23 characteristics that were retrieved from speech signals and characterise the voice measure as well as their interpretation. The Parkinson disease patient information attributes are highlighted in Figure 2.

Variable	Description
name	ASCII subject name and recording number
MDVP:Fo(Hz)	Average vocal fundamental frequency
MDVP:Fhi(Hz)	Maximum vocal fundamental frequency
MDVP:Flo(Hz)	Minimum vocal fundamental frequency
MDVP:Jitter(%)	measures the percentage variation in fundamental frequency
MDVP:Jitter(Abs)	the absolute value of jitter
MDVP:RAP	
MDVP:PPQ	
Jitter:DDP	Several measures of variation in fundamental frequency
MDVP:Shimmer	
MDVP:Shimmer(dB)	
Shimmer:APQ3	
Shimmer:APQ5	
MDVP:APQ	
himmer:DDA	Several measures of variation in amplitude
NHR	
HNR	Two measures of ratio of noise to tonal components in the voice
status	Health status of the subject (one) - Parkinson's, (zero) - healthy
RPDE, D2 -*	Two nonlinear dynamical complexity
DFA -*	Signal fractal scaling exponent
spread1,spread2,PPE	Three nonlinear measures of fundamental frequency variation

Figure 2: Attribute Information of Parkinson Disease Benchmark Dataset

The Parkinson disease dataset information is categorized into two parts namely training and testing. As shown in Figure 3, the 75% of PD diagnosis dataset are used for training and remaining 25% datasets are used for testing.

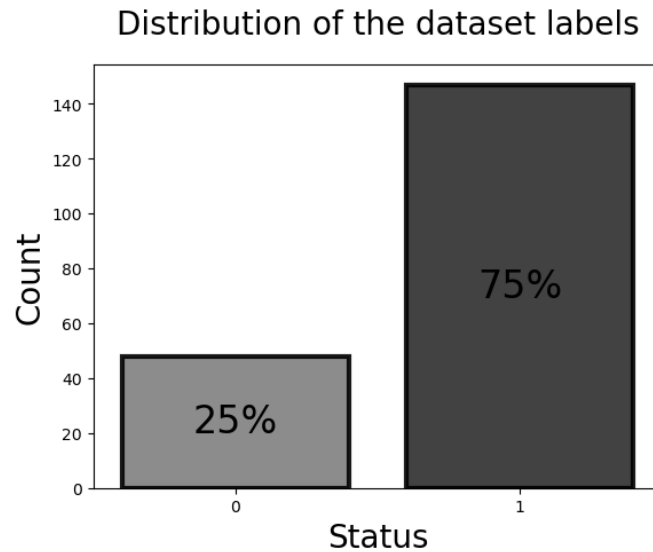


Figure 3: PD Dataset Split-Up

2. Pre-Processing Stage

In order to make sense of large amounts of data, processing is required. One of the most crucial actions to guarantee the effectiveness of future measures is data analytics. First, data must be generated by filling in missing values, eliminating outliers, and getting rid of duplicates [46]. Second, data must be validated to make sure it's full and consistent. In this article, we discovered that the dataset does not include duplicate values as the number of rows is equivalent to the number of unique column values. Except for the "status" feature, which is a binary categorical type, we also observe that all the features are continuous "numerical variables" kinds. If errors are found in the data processing processes, the corrective measures are determined by the kind and severity of the errors. Imputing missing values, dealing with outliers, eliminating duplicates, and examining and fixing data-validation errors are all examples of what this entails. The data's quality and completeness must be guaranteed in order to conduct trustworthy research. The each and every attributes of Parkinson disease dataset and minimum, maximum values ranges are represented in Figure 4.

AN INTEGRATIVE APPROACH TO DIAGNOSING
PARKINSON'S DISEASE USING ENSEMBLED MACHINE LEARNING MODELS

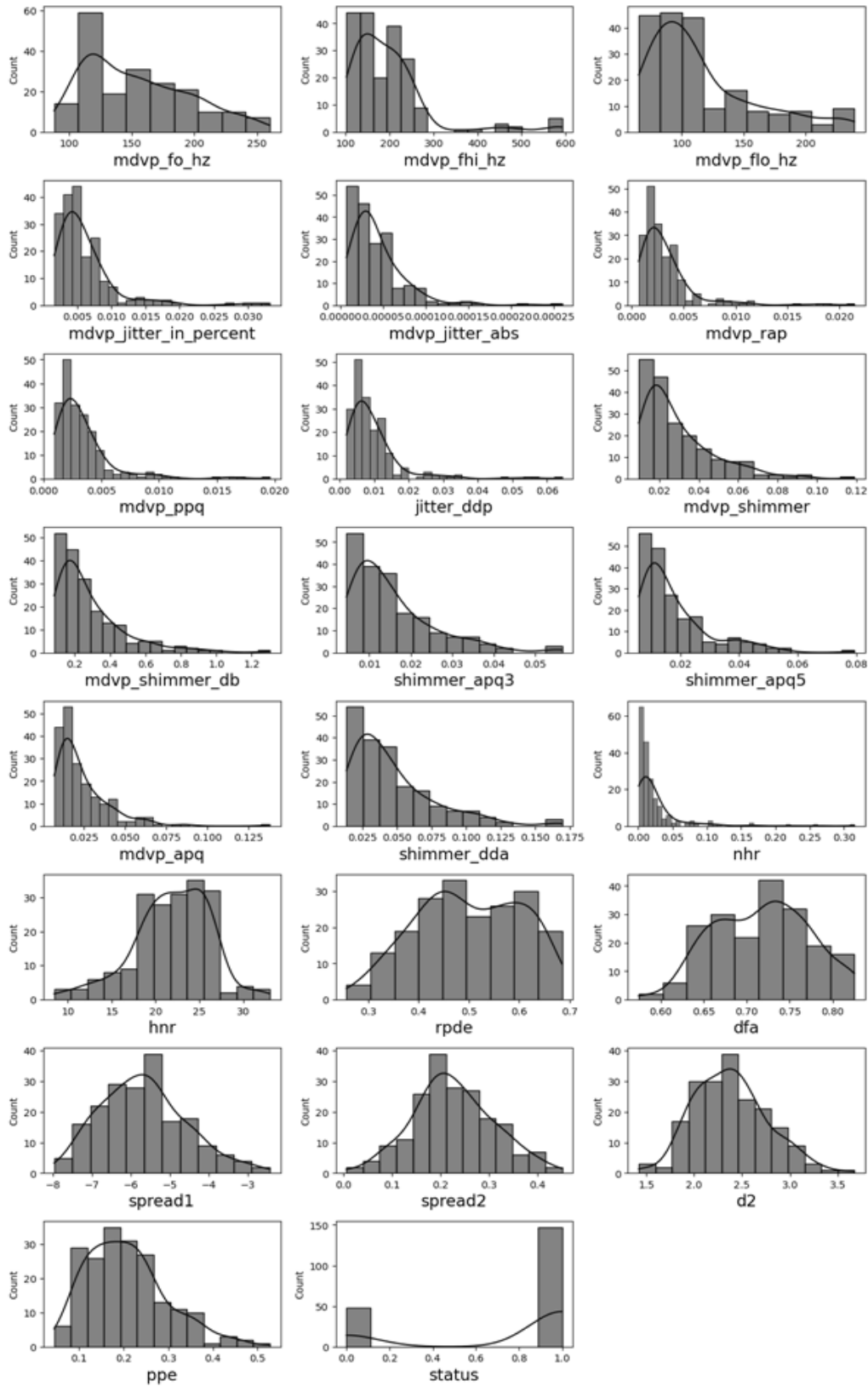


Figure 4: Maximum, Minimum Value Ranges of Parkinson Disease Dataset

3. Feature Extraction Stage

The scale invariant feature transform was initially presented by David Lowe [23]. It does this by searching through photographs for sites of interest and providing local descriptions that explain the surrounding neighbourhood. The initial phase of this approach is to locate extrema in the picture after it has been processed using the Difference of Gaussian (DoG). It achieves scale invariance by filtering the input picture at a variety of sizes and then gradually decreasing its sample size. The pixels are then compared to one another. The levels immediately adjacent to it, both lower and higher, are also evaluated. A pixel that is either the maximum or the lowest of all the pixels that are contiguous is a potential key point.

After then, a more in-depth analysis of the major topics is performed so that the "best" choices may be chosen. Each of the essential points has a stable foundation. Low contrast and precarious edge spots have been eliminated. After that, an orientation is assigned to every remaining key point. The technique takes use of gradient orientations that are close to the pixel. The gradient magnitudes are used as a weighing factor for the values. Feature vectors, also known as descriptors, are constructed using the produced points. The calculation employs the 16x16 neighbourhood surrounding the pixel. It is possible to compute the gradient magnitudes and orientations of the neighbourhood. Their values are weighted using a Gaussian distribution. In this circumstance, the orientation histograms are created for each sub-region separately. The result of this process is a vector containing 128 values (16 x 8).

4. Traditional Classification

Various conventional classification approaches have been employed for the diagnosis of Parkinson's disease, utilising a range of factors extracted from medical data. These features encompass clinical evaluations, imaging investigations, and demographic information. In this article, we have used five traditional classification techniques such as SVM, decision tree, k-nearest neighbors, logistic regression and naïve bayes classifiers.

5. Ensembled Classification

The term ensemble classification also refers to the process of "learning by committee." It's a powerful tool for boosting the reliability of categorization systems. The ensemble model [24] may take learners that are only marginally better than a random guess and combine them into powerful learners that can generate reliable predictions. More so than with base learners, an ensemble classifier has a superior capacity for generalisation. Because of its ability to be taught from existing data and then applied to new data, ensemble classifiers are often supervised learning algorithms. Generating the basis learning and then merging the base learners are the two main components of an ensemble. Training data is used to create a base learner, which is then fed into a base learning algorithm (such as a decision tree, neural network, or any other machine learning algorithm). Ensemble classifier approaches are already obtained improved success in actual job applications such as image identification, medical diagnosis and disease classification. Voting, bagging, boosting, and stack generalisation are the four mainstays of traditional ensemble techniques that we concentrate on here. When used to PD disease diagnosis, ensemble classifiers are most often employed to boost performance and increase the prediction rate.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Python and the Anaconda IDE were utilised in the creation of the suggested ensemble model, which incorporates AdaBoost, Bagging, soft voting, and weighted voting. The model was used to analyse the data from the PD disease dataset from UCI Machine Learning Repository by Max Little of Oxford University. The speech signal collection includes 195 biomedical sounds, of which there are 147 phonetics for persons with Parkinson's disease (PD) and 48 for healthy people. The training phase and the testing phase have both been separated from the dataset. The training phases consist of the training of sample feature in the ensemble model. During the testing phases, you will try to predict the outcome of test data by using the training datasets.

Experimental Analysis with Base Classifiers

Within this particular area, an examination has been conducted on the effectiveness of fundamental classifiers like Support Vector Machine (SVM), decision tree (DT), k- nearest neighbors (k-NN), logistic regression (LR) and naïve bayes (NB) classifiers. The mean values of accuracy, precision, recall, and F1-score for each base classifier were evaluated and shown in Table 2. The Support Vector Machine models had the highest level of performance accuracy, reaching 91.2%. The decision tree algorithm and kNN model achieved the second greatest level of performance, with an accuracy rate of nearly 88%. The naïve bayes classifier had the lowest classification performance among the five basic classifiers, achieving an accuracy rate of 81.2%.

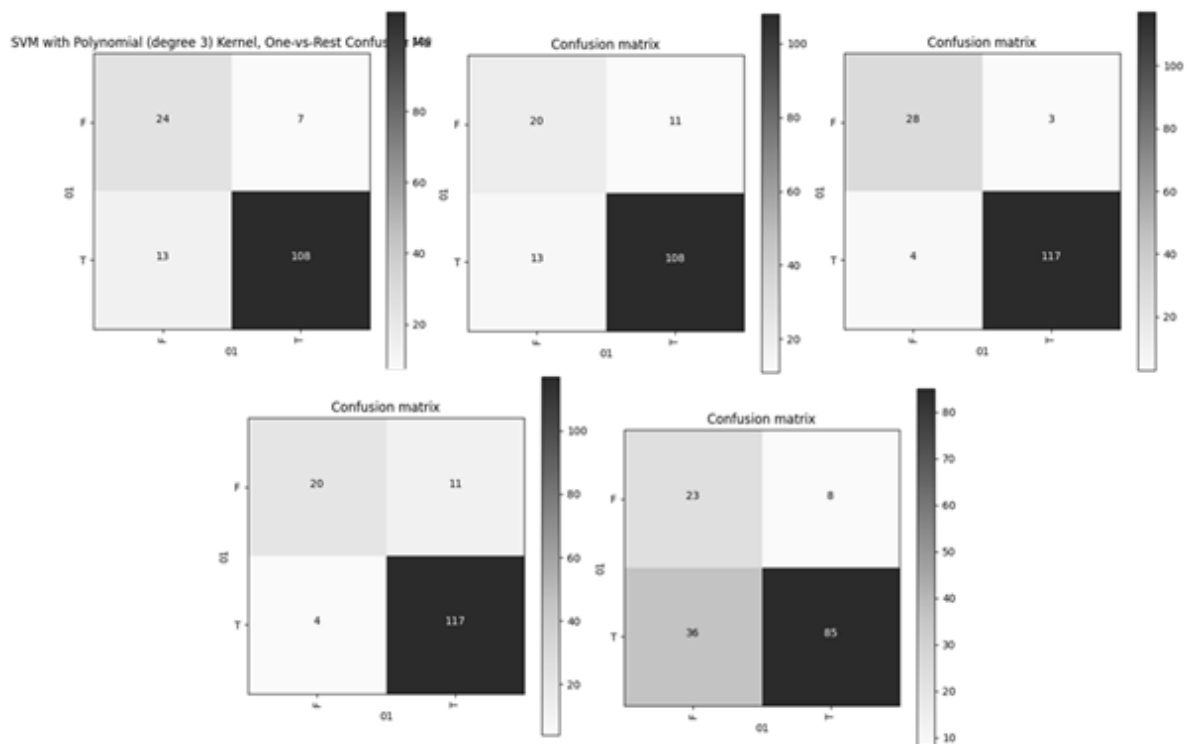


Figure 5: Confusion Matrix of SVM, Decision Tree, K-NN, Logistic Regression and Naïve Bayes Classifiers

Table 2: Performance Analysis of Base Classifiers

Sl. No.	Base Classifiers	Accuracy	Precision	Recall	F1-Score
1.	Support vector machine	91.2	90.4	90.8	90.6
2.	Decision Tree	88.4	88.1	88.5	88.3
3.	kNN	87.1	86	86	86
4.	Logistic Regression	84.7	84.6	84.1	84.3
5.	Naive Bayes	81.2	83	82	82.5

In order to optimise the efficacy of individual or basic classifiers, it is important to include them with additional classifiers. In the framework of our combined endeavours, we have effectively formulated a set of five discrete ensemble learning models.

- The bagging technique has been employed in the support vector machine classifier.
- The Adaboost approach has been implemented in the support vector machine classifier.
- In the context of the weighted voting technique, we have employed an ensemble approach consisting of three basic classifiers: Decision Tree, Multilayer Perceptron (MLP) and Support Vector Machine.
- In the end, the majority voting technique utilises an ensemble of basic classifiers, including Decision Tree, Multilayer Perceptron (MLP) and Support Vector Machine (SVM).

Table 3: Performance Analysis of Ensemble Classifiers

Sl. No.	Ensembled Classifiers	Accuracy	Precision	Recall	F1-Score
1.	SVM + AdaBoost	93	93.1	93.4	93.2
2.	SVM + Bagging	92.25	91.75	91.5	91.6
3.	MLP + DT + SVM + Majority Voting	93.50	94	94	93
4.	MLP + DT + SVM + Weighted Voting	95.7	95.1	95.4	95.3

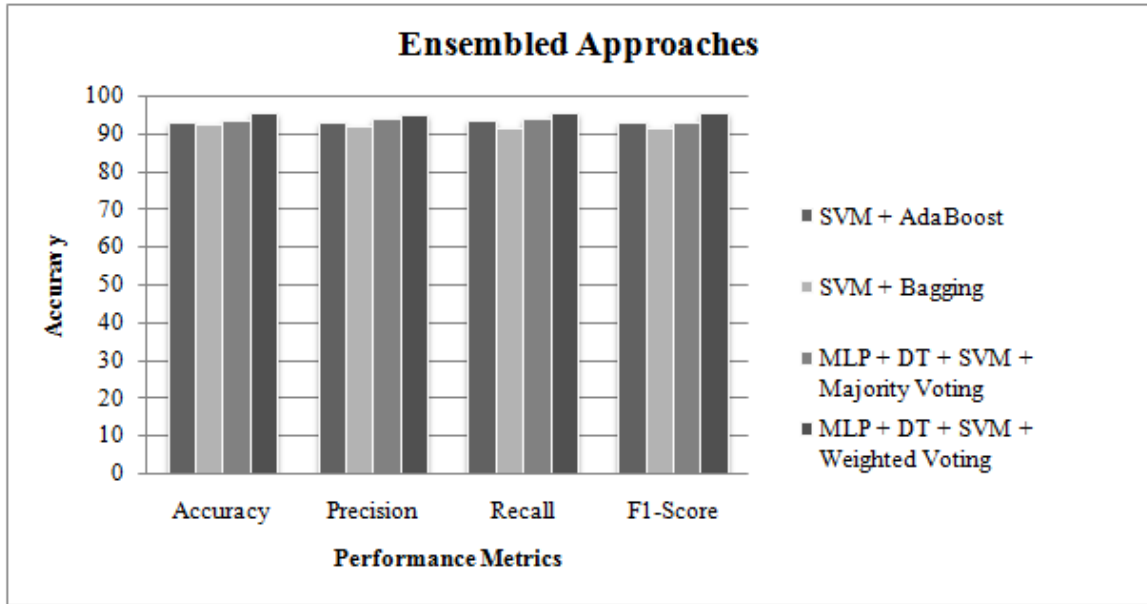


Figure 6: Performance Comparison with Proposed Ensembled Approaches

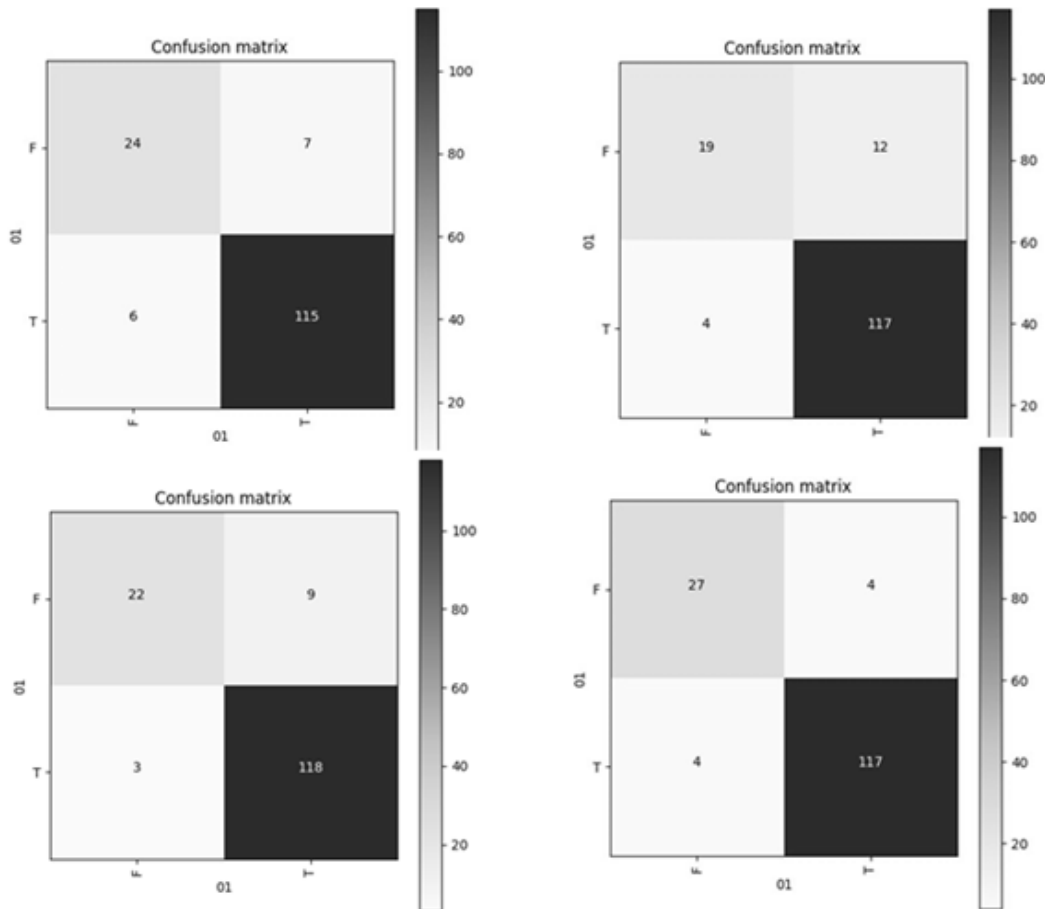


Figure 7: Confusion Matrix of Ensemble Adaboosting, Bagging, Majority Voting and Weighted Voting

Table 3 and Figure 6 present the performance of several ensembling strategies on the basis models, specifically SVM-Linear, decision tree, kNN, Logistic Regression and naïve bayes. The table suggests that the ensemble models exhibit superior performance, despite the somewhat lower performance of the individual base classifiers. Furthermore, it was observed that the implementation of the majority voting strategy using decision tree, MLP, SVM-Linear, yielded the best accuracy of 95.7%. Conversely, the bagging technique employed with Random Forest resulted in the lowest accuracy of 91.25%. However, it surpasses all five of the fundamental classifiers employed. The confusion matrix of the ensemble models is depicted in Figure 7.

V. CONCLUSIONS AND FUTURE ENHANCEMENTS

The classification and prediction of Parkinson disease in voice signal is a formidable challenge due to the inherent diversity in the varying of frequency belonging to the same category. In order to achieve favourable outcomes, a hybrid strategy known as ensembling has been implemented. Ensembling involves training numerous models instead of relying on a single model, and subsequently combining the predictions generated by these models. This study evaluates four distinct ensemble algorithms, including Adaboost, bagging, majority voting and weighted voting. Furthermore, it has been discovered that the majority voting technique demonstrates the highest performance among the suggested ensemble models. Conversely, boosting exhibits the lowest performance. However, it still outperforms the best individual base classifier models, namely support vector machine and naïve bayes.

To handle large dataset in the machine learning ensembled method is difficult, so in future we plan to implement the deep learning model and ensemble deep learning classifier. Also reduce the training time of future work we have planned to implement in the GPU environment.

REFERENCES

- [1] DeMaagd, G.; Philip, A. Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis. *Pharm. Ther.* 2015, 40, 504–532.
- [2] Rizek, P.; Kumar, N.; Jog, M.S. An update on the diagnosis and treatment of Parkinson disease. *CMAJ* 2016, 188, 1157–1165. [CrossRef] [PubMed]
- [3] Available online: <https://www.who.int/news-room/fact-sheets/detail/parkinson-disease> (accessed on 30 October 2022).
- [4] De Rijk, M.C.; Launer, L.J.; Berger, K.; Breteler, M.M.; Dartigues, J.F.; Baldereschi, M.; Fratiglioni, L.; Lobo, A.; Martinez-Lage, J.; Trenkwalder, C.; et al. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. *Neurologic Diseases in the Elderly Research Group. Neurology* 2000, 54 (Suppl. S5), S21–S23. [PubMed]
- [5] Cantürk, I.; Karabiber, F. A machine learning system for the diagnosis of Parkinson's disease from speech signals and its application to multiple speech signal types. *Arab. J. Sci. Eng.* 2016, 41, 5049–5059. [CrossRef]
- [6] Singh, N.; Pillay, V.; Choonara, Y.E. Advances in the treatment of Parkinson's disease. *Prog. Neurobiol.* 2007, 81, 29–44. [CrossRef]
- [7] Rana, A.; Rawat, A.S.; Bijalwan, A.; Bahuguna, H. Application of multi-layer (perceptron) artificial neural network in the diagnosis system: A systematic review. In *Proceedings of the 2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*, San Salvador, El Salvador, 22–24 August 2018; pp. 1–6.
- [8] Lakany, H. Extracting a diagnostic gait signature. *Pattern Recognit.* 2008, 41, 1627–1637. [CrossRef]

- [9] Figueiredo, J.; Santos, C.P.; Moreno, J.C. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. *Med. Eng. Phys.* 2018, 53, 1–12. [CrossRef]
- [10] Hazan, H.; Hilu, D.; Manevitz, L.; Ramig, L.O.; Sapir, S. Early diagnosis of Parkinson's disease via machine learning on speech data. In *Proceedings of the 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel, Eilat, Israel 14–17 November 2012*; pp. 1–4. [CrossRef]
- [11] Karan, B.; Sahu, S.S.; Mahto, K. Parkinson disease prediction using intrinsic mode function based features from speech signal. *Biocybern. Biomed. Eng.* 2019, 40, 249–264. [CrossRef]
- [12] Frid, A.; Safra, E.J.; Hazan, H.; Lokey, L.L.; Hilu, D.; Manevitz, L.; Ramig, L.O.; Sapir, S. Computational diagnosis of Parkinson's Disease directly from natural speech using machine learning techniques. In *Proceedings of the 2014 IEEE International Conference on Software Science, Technology and Engineering, Washington, DC, USA, 11–12 June 2014*; pp. 50–53.
- [13] Rawat, A.S.; Rana, A.; Kumar, A.; Bagwari, A. Application of multi layer artificial neural network in the diagnosis system: A systematic review. *IAES Int. J. Artif. Intell.* 2018, 7, 138.
- [14] Das, R. A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Syst. Appl.* 2010, 37, 1568–1572.
- [15] Al-Sarem, M., Saeed, F., Boulila, W., Emara, A.H., Al-Mohaimed, M., Errais, M.: Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson's disease. *Advances on Smart and Soft Computing. AISC*, vol. 1188, pp. 189–199. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-6048-4_17
- [16] Wroge, T.J., Ozkanca, Y., Demiroglu, C., Si, D., Atkins, D.C., Ghomi, R.H.: Parkinson's disease diagnosis using machine learning and voice. In: *2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–7. IEEE (2018)
- [17] Wan, K.R., Maszczyk, T., See, A.A.Q., Dauwels, J., King, N.K.K.: A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease. *Clin. Neurophysiol.* 130(1), 145–154 (2019)
- [18] Cavallo, F., Moschetti, A., Esposito, D., Marenmani, C., Rovini, E.: Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning. *Parkinsonism Related Disord.* 63, 111–116 (2019)
- [19] Almeida, J.S., et al.: Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recogn. Lett.* 125, 55–62 (2019)
- [20] Parisi, L., RaviChandran, N., Manaog, M.L.: Feature-driven machine learning to improve early diagnosis of Parkinson's disease. *Expert Syst. Appl.* 110, 182–190 (2018)
- [21] Kotsavasiloglou, C., Kostikis, N., Hristu-Varsakelis, D., Arnaoutoglou, M.: Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomed. Signal Process. Control* 31, 174–180 (2017)
- [22] Lichman, M. UCI Machine Learning Repository; University of California, School of Information and Computer Science: Irvine, CA, USA; Available online: <http://archive.ics.uci.edu/ml> (accessed on 25 September 2022).
- [23] David G. Lowe. Object recognition from local scale-invariant features. In *International Conference on Computer Vision*, 1999.
- [24] S. Khalid, T. Khalil, and S. Nasreen, "A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning." *International conference on Science and Information*, pp.372–378, 2014.