

ADVANCEMENTS IN ELECTROENCEPHALOGRAPHY AND COGNITIVE PROFILING FOR PARKINSON'S DISEASE

AANCHAL SHARMA¹, ANU GUPTA¹ AND SUKESHA SHARMA²

¹RESEARCH SCHOLAR AND ¹PROFESSOR DCSA, PANJAB UNIVERSITY, CHANDIGARH, INDIA, EMAIL:
aanchal6sharma@gmail.com

²ASSOCIATE PROFESSOR UIET, PANJAB UNIVERSITY, CHANDIGARH, INDIA

ABSTRACT

Background: Parkinson's disease (PD) is a neurological defect that causes both non-motor (brain-related) and motor (physical) symptoms. As PD progresses, brain function (cognition) may diminish in some cases. Over recent times, quantitative electroencephalographic (QEEG) has emerged as a valuable tool for investigating the growth of cognition, psychological variables and psychopathological illnesses (mental disorders).

Objective: The principal objective of this study is to convey knowledge about analysis methodologies, novel approaches and practical information to those with little expertise, with future directions.

Methods: This paper covers the EEG preprocessing steps with methods, analysis techniques, transformation methods and classification details from the research conducted between 2006 to 2024. There exists no study categorizing the signal processing techniques based on the different domains. Therefore, to fill this gap, this paper focuses particularly on the analysis categories that are time-frequency domain (TFD), frequency domain (FD), time domain (TD), graph theory-based, network dynamics, functional connectivity and connectivity analysis techniques with their strengths and limitations. Additionally, an experiment is carried out to investigate the ability of EEG to categorize cognitive processes and the relevancy of machine learning approaches in determining the best classifier to handle EEG data.

Conclusion: Analysis of 59 studies revealed that non-neural machine learning techniques are the most frequently employed. Additionally, the highest accuracy rates from recent studies are identified and the usage percentages for each model are also calculated. Therefore, the purpose of this study is to provide researchers with an initial base for future research on cognition in PD.

Keywords: EEG, Machine Learning, QEEG, Cognitive Profiling, DBS, PD, ICA, GNN

[1] INTRODUCTION

PD is the most prevalent source of dementia and mobility (movement) impairment in elderly people. James Parkinson, who initially defined the condition, has come to represent the disorder [1]. In the latter stages of PD, mental decline is a prevalent symptom. Problems with attention, memory and visual-spatial skills are some of the other, brain related symptoms. The dopaminergic deficit is the cause of PD and this deficiency is the consequence of neuronal death in the brain region known as the Substantia Nigra [2]. Patients who lived beyond 10 years following their diagnosis are predicted to acquire dementia in 75% of cases. Detecting cognitive difficulties in PD patients by the use of QEEG has shown encouraging results. Medications that boost dopamine levels, such as levodopa, are used to treat the disease's symptoms. The benefits of levodopa for treating PD-related motor symptoms are only transient and patients may still have motor problems during "OFF" time when the drug is no longer effective. There is presently no approved disease-modifying treatment for PD-related cognitive decline [3]. Potential therapies for Parkinson's disease symptoms and progression include neuro-modulation, physical activity and cognitive rehabilitation [4]. In cognitive rehabilitation treatment, therapists and psychologists collaborate with medical professionals to develop a strategy to help patients restore as much cognitive function as feasible. Cognitive functions include multitasking speed, concentration power, memory, processing speed, etc. A Neuro-modulator refers to a substance that alters nervous system function in reply to an outside stimulus, such as electrical pulses [5]. Motor symptoms respond better to neuro-modulators by Deep Brain Stimulation (DBS). DBS uses a stimulator (a surgically implanted small device) that sends signals to the targeted area of the brain which helps to improve motor functions. Despite the positive benefits on motor performance, some decline in cognitive and neuropsychiatric areas, such as cognitive impairment [6], depression [7], speech intelligibility [8], etc., might be found after surgery. Therefore, it becomes very necessary to know about the cognitive status of the candidates before they go through the DBS screening [9]. Dementia (PDD), Mild Cognitive Impairment (MCI) and Subjective Cognitive Decline (SCD) are all possible forms of cognitive disorder in PD patients. Patients with SCD typically experience only a slight decrease in cognitive ability and their performance on neurocognitive tests falls within the normal range [10]. In contrast, the cognitive decline shown in PD-MCI is gradual but constant. Visual-spatial skills, mental processing speed, memory, planning & organization, language skills and social cognition are the

six distinct subfields of cognition. PD-MCI is classified as either various domain or just one domain depending on the number of affected cognitive regions [11]. Memory, concentration, executive function and visuospatial abilities are required for a diagnosis of PD-D. People with PD-D often experience hallucinations, depression and a lack of motivation [12]. PD and PD-MCI are defined using different criteria in the literature. The MDS-PD (Movement Disorder Society Task Force) is the most often utilized criterion to diagnose PD-MCI [13]. After following them for three years, researchers found that many PD patients likely retain some degree of MCI [14]. Even though 60% of those with MCI would develop PD-D within 4-12 years, according to certain research [15]. Numerous methods have been put out in the literature to evaluate cognitive function in PD. PD patients may be evaluated for their cognitive abilities using a variety of tests, including the MoCA (Montreal Cognitive Test), UPDRS (Unified PD Rating Scale) and MMSE (Mini Mental State Examination) scores. Multiple criteria are used in the literature to assess patients with PD and their cognitive abilities [16] [17].

[1.1] Electroencephalogram (EEG)

An EEG can capture the brain's electrical activity. In 1875, an English doctor named Richard Caton found that electrical currents could be detected from the exposed brains of rabbits using simple radio equipment. Later, a German neurologist by the name of Hans Berger confirmed this finding. The brain's electrical signal is a miniature representation of the mind's operations. Clinically, this is of interest in the study of epilepsy, brain injury, strokes, tumors and behavioral anomalies, etc. Brain waves may be broken down into five distinct categories: gamma, alpha, theta, beta and delta. [18] (displayed in Table 1).

Table 1: Details of EEG frequency bands				
Frequency Bands	Location	Frequency Range	Useful area	Nature
Delta (δ)	Frontal Lobe	0 to 4 Hz.	Measures deep sleep	high waves Amplitude, slow
Theta (θ)	Central, Temporal	4 to 8 Hz.	Measures the stages of sleepiness that are most relevant to meditative states and daily life	high waves Amplitude, slow
Alpha (α).	Frontal, Occipital	8 to 12 Hz.	Used to evaluate one's mental state, focus or concentration	low waves Amplitude, fast
Beta (β)	Central, Frontal	12 Hz. to 30	Helpful for gauging the extent of brain injury	low waves Amplitude, fast

An in-depth study of the brain is made possible by collecting EEG data over a wide range of frequencies. Quantitative electroencephalographic analysis (QEEG) is often known as "brain mapping." QEEG analysis goes beyond visual EEG interpretation, which may help us learn more about the brain and how it works [19]. Trained experts may utilize the EEG and the accompanying QEEG data to assess brain function and monitor the effects of therapies like neuro- feedback and medication. The first step in the EEG analysis is signal acquisition. It is the process of capturing and storing electrical activity produced by the brain using an electroencephalogram (EEG). Several electrodes are applied to the scalp in order to identify and quantify the electrical impulses produced by brain neurons. These electrodes (sensors) are linked to an EEG machine or device, which amplifies & filters the electrical impulses before recording them for examination. Pre- & post-processing of EEG data are crucial steps after acquiring signals. The analysis of signal can be performed using various methods as displayed in Fig. 1. Preprocessing is essential to get a clean signal for feature extraction. Independent Component (IC) Analysis and Principle Component (PC) Analysis are often used to eliminate artifacts in ECGs (Electrocardiograms) and EOGs (Electrooculograms). Instead of this, there are many preprocessing pipeline methods, which consist of the number of steps that get automatically executed.

There are benefits and limitations to each of these approaches. The researcher can use different preprocessing pipeline methods according to his/her requirements.

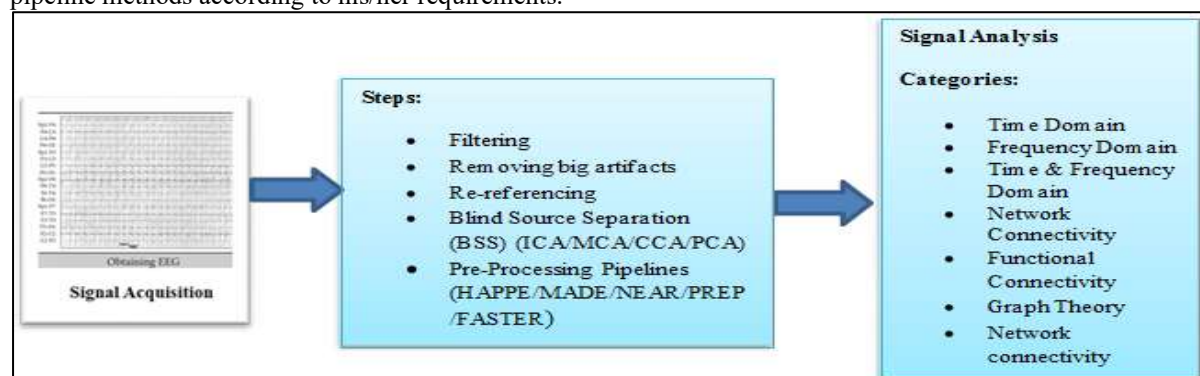


Fig. 1: Main steps of the processing

The phases of cognition in PD have been identified in this work by a review of QEEG. This study is structured as follows: section 2 includes an extensive literature review that covers all EEG analysis approaches and transformation methods with the findings and conclusions of the previous studies using these analysis techniques. This section covers difficulties and restrictions in the related existing research. A Machine Learning (ML) model comparison for anticipating cognitive phases in PD is discussed in Section 3. Section 4 discusses the results, along with some observations and discussions for the future. Section 5 presents the findings and conclusions.

[2] RELATED WORK

Various EEG analysis techniques are available in the literature to analyze the EEG signals of PD patients. These include individual alpha frequency-based, spectral, occipital peak frequency analyses, phase lag index, weighted & directional phase lag index and graph theory-based analysis. This paper also emphasizes the importance of analysis methods. Furthermore, these analysis methods can help researchers to develop new therapeutic approaches to improve their results. EEG signals are typically recorded in the time domain. However, in certain applications, it is necessary to analyze the signals in the FD. To achieve this, signal conversion techniques are used to convert the EEG signals from TD to FD or TFD. These conversions provide insights into the frequency content and time dynamics of the signals.

Based on the literature, IAF (Individual Alpha Frequency) has been examined for frequency domain analysis [20] [21] [22]. It refers to the frequency range of 8 to 12 Hz at which an individual's alpha brainwaves oscillate. It is observed that healthy people will have higher mean IAF as compared to PD-MCI patients. Similarly, spectral power distribution among a wide frequency range gives information about various levels of cognition in Parkinson's disease [23] [24]. High theta band and low alpha band spectral power have been observed for PDMCI or PDD patients as compared to non-demented patients of PD. This suggests that people with PDD have more spectral power in the theta band (4.0-8Hz) than those with PD-NC or healthy controls. In PD, not only spectral power but also functional interconnectedness (associations between several areas of the brain) has been investigated. The alpha band correlation is often lower in people with cognitive impairment. In one of the studies, to automate the separate steps of preprocessing, feature extraction and classification, a pipeline technique has been used [25]. It is analyzed that the pipeline model's accuracy for the intermediate stage is low (falls in between good or bad cognition) prediction because it has a very large number of features. With such a large number of features, the problem of overfitting occurs which leads to high training accuracy and low-test accuracy. A pipeline is a sequential method that involves breaking down a difficult activity or procedure into several stages or steps, where each stage performs a particular task and passes the outcome to the next stage as input. Each stage of the pipeline runs separately and concentrates on one specific aspect of the entire procedure. Various issues observed from the literature are the low accuracy of models, data variability among subjects and the non-linear & complex nature of EEG signals. Therefore, analysis methods are a crucial component of EEG in Parkinson's disease to cover all these issues.

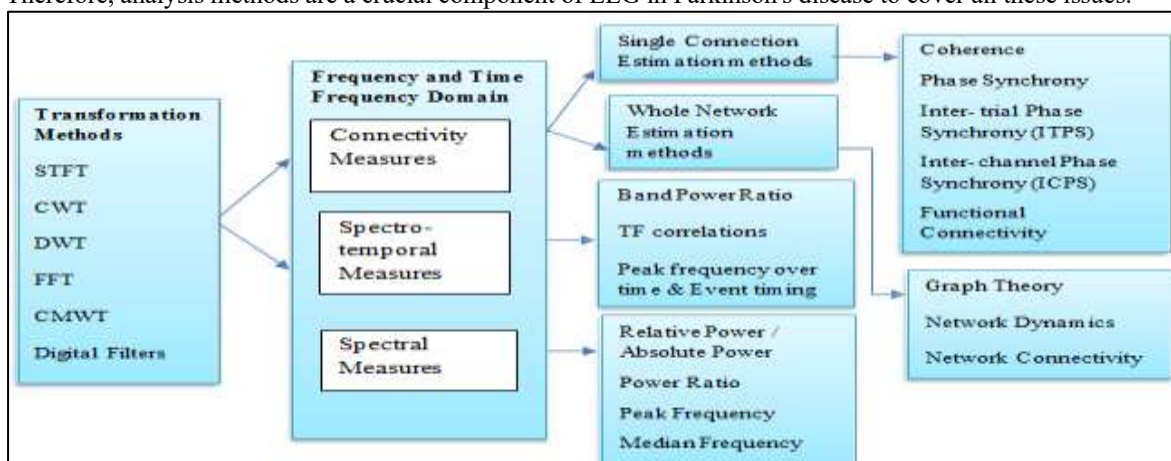


Fig. 2: Spectral, Spectro-temporal and connectivity measures with transformation methods

[2.1] Quantitative analysis methods to classify cognition

To classify cognition, analysis of EEG signals can be done using Quantitative electroencephalography (QEEG). A QEEG procedure involves several stages. The initial part of the process is acquiring an EEG signal, which may be done using several different types of EEG equipment and electrode setups. Preprocessing is the second step and it entails removing artifacts from the EEG such as muscle activity, tiredness, eye blinks, pulse rate and so on. As part of the preprocessing phase, clean EEG segments are selected for further analysis. Finally, linear or non-linear mathematical techniques can be used to clean the EEG signal to extract a parameter that is required for further processing. When EEG signal is assumed as a steady state signal then linear techniques can be used for the analysis. In reality, the EEG signals exhibit irregularity and erratic nature. Therefore, some non-linear approaches are also available for analysis. Linear techniques that are used to remove the artifacts are band-pass filters, notch filters and adaptive filters. Whereas Blind-Source Separation (BSS) methods such as ICA, Multichannel Convolutional Autoencoder (MCA) and PCA are non-linear techniques (Fig. 1). ICA divides mixed

EEG signals into statistically independent components and isolates brain-related signals to remove artifacts. The MCA method is deep learning (DL) based method which discovers the connection between mixed EEG signals and their underlying sources, enabling efficient artifact removal. PCA decreases the dimensionality of EEG data, by separating the orthogonal components from EEG data. These all methods are effective in artifact detection and removal. After pre-processing, the next step is feature extraction, information on EEG signals is given in the time domain, but those signals do not give the exact frequency information. So, those signals are transformed into the frequency domain using different transformation methods (Fig. 2) like the Fast Fourier Transform (FFT), Welch method, Short-Time Fourier Transform (STFT), Multi-taper method, Discrete and Continuous Wavelet Transform (DWT & CWT) and Complex Morlet Wavelet Transform (CMWT). These methods give detailed information about the frequency bands with time and spectral power of each band in the whole spectrum. After domain transformation, EEG characteristics can be analyzed by different types of measures i.e. connectivity measures, spectro-temporal measures and spectral measures as shown in Fig. 2. Spectral measurements are used to measure the distribution of power across each band in the EEG signals. They include relative power, absolute power, median frequency and power ratio, etc.

The spectro-temporal measures are used to give the time-frequency features such as energy distribution, band power ratio, TF correlations, peak frequency over time and event timing, etc. Connectivity measures, instead, concentrate on the relationships between various brain regions. whereas, focus on the interactions between different brain regions. There are two types of connectivity measures: single network estimation and whole network estimation. Single network estimate is used to investigate the functional connectivity and communication patterns between specific brain area pairs. Single network estimation measures include coherence, phase synchrony, cross-correlation, Inter-trial & Inter-channel Phase Synchrony (ITPS & ICPS) and Functional Connectivity, etc. On the other hand, whole network estimation analyzes the connectivity patterns across the entire brain network. It provides a comprehensive understanding of how different brain regions connect and interact with one another. Whole network estimation methods involve techniques such as graph theory analysis, network connectivity analysis and network dynamics analysis, etc. These methods help to understand the overall organization, efficiency and integration of the brain network, which further helps in detecting the cognitive impairment of PD patients.

[2.1.1] Transformation methods: Several methods exist for generating EEG sub-bands, including FFT, digital filters and STFT & wavelet transform. Table 2 lists many common transformations used in EEG signal analysis. Converting data between different domains, such as from the TD to FD and back, is a popular use of FFT [26]. The FFT technique works well with stationary linear inputs but is very sensitive to background noise. Unlike other digital signal processing approaches, FFT operates poorly with non-stationary signals because it cannot pick up on small changes in the signal [27]. As a result, there is confusion over when certain frequency-domain features are recorded. Thus, FFT is not preferred in many modern EEG-based applications, such as monitoring driver sleepiness or identifying sleep disorders, etc. In addition to separating the data into distinct EEG frequency bands, the FFT may be utilized to extract statistical metrics [28]. To analyze non-stationary data, STFT may be used, a window-based variation of FFT. STFT operates by performing the FT on each frame of the signal. Instead of analyzing the complete signal, this permits the frequency content to be studied over time. The resultant STFT may be visualized as a spectrogram, a two-dimensional depiction of the signal's frequency content against time. The resolution of the frequency analysis in the STFT is limited by the length of the analysis window, which is a downside of the STFT. It will be difficult to see the time-varying behavior of particular frequency components if the analysis window is too small, this results in high frequency resolution but low temporal resolution. However, if the analysis window is too big, both time and frequency resolution will be inadequate, making it impossible to distinguish between the various frequency components. Therefore, choosing an appropriate analysis window is important for achieving a fair balance between time & frequency resolution in STFT analysis. Additionally, STFT may not be well-suited for analyzing non-stationary signals with rapidly changing frequency content, as it assumes the signal is stationary over the length of the analysis window [29] [30]. The easiest method to extract the sub-bands from digital filters is to use an FIR filter. Among the two approaches, the frequency technique and the windowing methods of FIR filters, the latter is superior for EEG analysis. FIR stands for Finite Impulse Response, which is a digital filter frequently employed in applications for signal processing. FIR filters are a class of linear filters, meaning that their output is a linear arrangement of their input signal samples. One of the advantages of FIR filters is that they have a linear-phase response, this implies they add a uniform delay to the whole signal at all frequencies. This may be helpful in fields like audio and picture processing where it's crucial to maintain the signal's phase information.

Non-stationary or time-varying signals in the EEG may be transformed using a time & frequency analysis approach called as a CWT. It gives a time & frequency representation of the EEG signal by decomposing it into a sequence of wavelets with varying scales. In contrast to CWT, which utilizes a continuous range of wavelets, DWT uses a sequence of discrete wavelets. Using a number of coefficients, DWT separates the EEG data into a range of frequencies. The energy or power of the EEG signal at different frequencies may be analyzed with the help of these coefficients. CWT and DWT methods are used in feature extraction techniques. However, the optimal method for analyzing an EEG signal relies on the task at hand and the nature of the signal itself [31].

Table 2: Different transformation methods for EEG signal analysis			
Transformation Methods		Strengths	Limitations
Wavelet Transform	CWT (Continuous Wavelet Transform)	a) Contains continuous parameters like time & scale b) best for extracting small features	a) parameters change causes inefficiency and duplication. b) high delay, limited to continuous data
	DWT (Discrete Wavelet Transform)	a) fast computation power b) supports both orthogonal and biorthogonal.	a) to choose decomposition levels b) selection of mother wavelet function is not clear
	CMWT (Complex Morlet Wavelet)	a) Goodtime-frequency resolution b) Robust to noise c) Capture both amplitude and information of EEG signals	a) Need expertise as parameters should be selected Carefully b) Complex computations
Digital Filters	IIR		b) limited to linear analysis
	FIR	Frequency	a) intuitive operation, much like FFT b) It can tolerate loads of any size response
		Window	a) easy in comparison to other approaches b) comfort of usage c) ability to adapt designs
Fourier Transform	FFT (Fast Fourier Transform)		a) gives the best results for stationary EEG signals b) fast Computations c) used to extract statistical features
	STFT (Short Time Fourier transform)		a) gives good performance if EEG signals are non-stationary b) fixed size window c) wide window consumes more time d) narrow window leads poor frequency resolution

[2.1.2] Analysis techniques: There are several categories for analysis of EEG: TD analysis, FD analysis, TF analysis, Phase Synchrony, Functional Connectivity, Network Dynamics, Network Connectivity and Graph Theory analysis. EEG signals are recorded in the TD, therefore, there are different methods available to convert the TD signal to the FD and TF domain and vice versa (As shown in Fig. 2). These approaches are used to transform and analyze EEG signals. Each analysis method is addressed in detail in the following subsections.

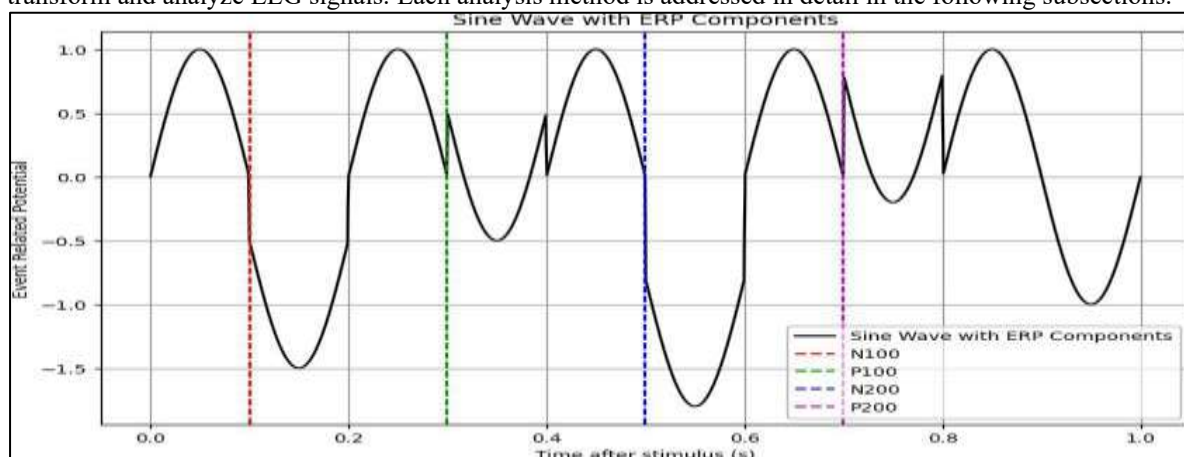


Fig. 3: Components of ERP waveform in time domain analysis

[2.1.2.1] TD (Time domain) analysis

TD or Time domain analysis can help to tell how a signal changes over time, there is no any direct information about frequency. Each data point in the signal represents a particular moment in time. It consists of properties like signal strength (amplitude), duration, shape and temporal interactions between distinct portions of the signal. The TD techniques are used to calculate the Event-Related Potential (ERP), measure peak amplitudes, detect event onset & offset times and analyze the signal shapes. ERPs are the electrical brain responses that occurs in response to certain events or stimuli. The brain's reaction to specific stimulus or event, related to cognition or motor is measured by ERP. It's an increase or decrease in voltage that results from brain activity in response to a stimulus [32]. A stimulus refers to an event or a condition that is presented to the subject during an experiment or recording session. Neurological and behavioral disorders, such as neurodegenerative illnesses, may be studied using the ERPs method [33]. Time domain analysis of EEG data reveals four major components of ERP waveform: The N100, N200, P200 and P300 (shown in Fig. 3). The N100 component is a negative waveform

deflection commonly seen about 100 ms after a sensory input has been presented. In addition to being characterized by a negative deflection in the waveform, the N200 component appears roughly 200 ms after the commencement of the stimulus. Attentional control, working memory and response inhibition are all assumed to be reflected in the cited study [34]. Around 200 milliseconds after a stimulus is presented, a positive deflection known as the P200 component appears. A positive deflection known as the P300 component happens 300–500 ms after a stimulus is given. Attention, working memory and decision-making are only a few of the cognitive processes that have been linked to it [35]. Researchers may learn about the underlying brain mechanisms of these processes by analyzing the temporal evolution and amplitude of these components. Peak detection, zero-crossing rate, waveform slope and energy calculations are a few examples of the features that are extracted from the time domain signals. Statistical metrics like median, mean, variance, STD (standard deviation) and skewness are computed to measure the central tendency, variability and distribution of the signal in the TD.

[2.1.2.2] FD (Frequency domain) analysis

The FD (frequency domain) is employed to examine the frequency component in the EEG signal. As mentioned earlier, alpha, delta, beta and theta have various ranges of frequencies. The FT of the EEG signal converts the signal from the TD to the FD to perform the spectral analysis. Spectral analysis is a technique for examining the frequency components by dividing a signal into its constituent frequency parts that can give vital information about the underlying mechanisms in the brain. The resultant power spectrum reveals the power distribution (power spectral density) across each frequency band. Many distinct mental states and cognitive operations have been linked to certain frequency ranges in the electromagnetic spectrum.

For instance, the alpha (8-12 Hz) band is connected with wakeful relaxation, whereas the beta (12-30 Hz) band is associated with alertness and focused attention and the delta (0.5-4.0 Hz) & theta (4.0-8 Hz) bands are generally associated with deep sleep and relaxation. It may also be used to examine how brain activity evolves over time during the execution of a mental task or in reaction to sensory input. Different research will probably utilize somewhat different cutoff frequencies. For a more in-depth analysis, the bands can be further divided into sub-bands like alpha 1 (8-10 Hz) and alpha 2 (20-30 Hz), etc. On the basis of the specific research question and the type of EEG data, some researchers use bi-spectral analysis when they observe that linear spectral analysis techniques (FTT, welch method, wavelet transform, stft method, etc.) are insufficient for capturing the nonlinear interactions in the data. The bi-spectral analysis examines the relationships between multiple frequency components of a signal, which can be useful in understanding the complicated systems. This analysis is performed by calculating the bicoherence or related measures [39]. Some previous studies that are based on spectral power analysis to classify the different cognition categories are tabulated in Table 3 [36-40]. The number of the participants, criteria used for cognition classification, regions used for analysis and the main findings using spectral analysis methods are also included in the table. This tabular data is useful for researchers aiming to identify cognition in PD because it provides insights into the analysis criteria and the findings of the previous papers.

Table 3: Studies related to spectral power (frequency-based) analysis using EEG data [23] [36-40]

Abbreviations employed in Table 3: PDD = PD with dementia; HC = Healthy controls; UPDRS = Unified PD Rating Score; DSM-IV = Diagnostic & Statistical Manual for Mental dysfunction; PD-N = PD normal control; AD = Alzheimer's disease; MMSE = Mini-Mental Score Examination; PD-MCI = PD with cognitive mild impairment.				
Author	Participants	Regions	Criteria used for classification	Main findings
Bosboom et al. (2006)	HC = 13, PD-NC=13, PD-D = 13	frontal, temporal, parietal and central, occipital	- EEG signals with a resting state and eyes closed - PD patients without dementia had an MMSE score of greater than 28. Patients with dementia had MMSE ≤ 24	- PD-NC in comparison to HC, relative theta power enhanced whereas beta power dropped. - The PDD patients' gamma strength was lower than that of the PD-NC group when evaluated. - PDD showed a drop in alpha & beta power and a rise in delta & theta power.
Ponsen et al. (2012)	PD-NC-13, PD-D-13	34 regions across the brain	- PD-NC patients with a MMSE score of >28. - PD-MCI with DSM-IV criteria for dementia and had a MMSE <24	- Individuals with PDD had more delta & theta strength than those with PD. - The PDD patients' frontal and parietal-temporal-occipital areas have lower alpha and beta power.
Fonseca et al. (2013)	AD = 38, PDD = 12, PD-NC = 31, HC = 37	Frontal temporal and occipital left-right	- closed-eyes resting posture EEG. - AD followed the DSM-IV standards.	- PDD has the highest delta and theta powers while PD-CN had the lowest. - PDD showed the highest alpha and beta frontal hemisphere coherence, while AD showed the lowest.

[2.1.2.3] Time & frequency-based analysis

Time and frequency (TF) analysis is concerned with how the amplitude and power of brain waves changes over time at specific frequency bands. This method breaks down MEG or EEG signals into their individual bands of

frequencies (alpha, beta, delta, theta), allowing researchers to identify the neural oscillations that are relevant for cognitive processing. This analysis can help the researchers to answer the questions like, "When does a particular frequency band become more active or suppressed during a cognitive task?" It is a helpful instrument for evaluating the characteristics of brain responses in terms of duration and frequency. The effective method used in TF is CWT with morlet wavelets, which may improve the clarity of this representation. The resulting spectrogram may appear more or less blurry, particularly for very low and higher frequencies. If the researchers do not have the exact prior knowledge of the frequency patterns they are interested in, then the spectrogram may appear blurry in terms of resolution [41, 42].

[2.1.2.4] Phase Synchrony and Connectivity Analysis

The connectivity analysis methods are used to give information on how the brain's many regions are connected and coordinate their functions with each other. TF analysis gives the phase information that can be used as a base to calculate the measures of phase synchrony in EEG data. Phase synchrony (phase coherence or phase locking) describes the degree of synchronization (coordination) or regularity in the timing of cyclical patterns of neural activity between different brain regions [41].

- **Functional connections** can be determined by investigating patterns of connectivity between various brain regions. The researchers can identify PD patients with different cognition on the basis of connectivity alterations in the brain. The term "functional connections" is coined to describe the relationship among the electrical activity recorded by several sets of electrodes in the same region or different regions of the brain [43, 44]. Coherence, phase lag index (PLI), weighted phase lag index (wPLI) & directed phase lag index (dPLI) are all examples of connectivity metrics that may be used to determine how well different regions of the brain are synchronized with one another [44-46]. Functional connectivity comes under the category of single connection estimate that investigates the connectivity or functional relationship between two brain areas or within a single brain region.

- **PLI** measures the consistency of phase relationships, particularly in the context of brain oscillations. It gives the output as 0 and 1. A PLI value of 0 implies that the two signals are leading or lagging each other equally and there is no consistent phase relationship. A value greater than zero implies that there is an imbalance in the likelihood of one signal leading or trailing the other. A value of 1 indicates that one signal consistently leads or lags the other signal. A greater PLI value indicates better phase synchronization, which can indicate a strong functional connection between brain areas.

$$PLI = |E[sgn(I(Xij))]|$$

In this equation, I is used to represent an imaginary component, sgn is used for sign, E is used to represent expected (mean) value and Xij is cross-spectral density between two signals i and j .

- **dPLI** is a PLI extension that provides directional knowledge regarding the phase connections of two signals. It can identify phase synchrony and determine which signal is moving ahead or lagging in direction. The null connectivity for dPLI is indicated by 0.5 (0,0.5,1), whereas the null connectivity for PLI and wPLI is 0 (0,1).

$$dPLI = E[H(I(Xij))]$$

In the given equation, I is used to represent an imaginary component, H is used for the Heviside step function, E is used to represent expected (average) value and Xij is cross-spectral density between two signals i and j .

- **wPLI** is another PLI form, which tries to improve the sensitivity to the actual phase synchronization. It helps to diminish the volume conduction effect. Volume conductivity effect is a situation when the electrical activity of one electrode spreads out to other electrodes and therefore, false electrical activity is recorded at several electrodes. This might produce inaccurate connection estimates. The wPLI solves this problem by allocating weights to the phase differences according to their dimensions. Because volume conduction is more likely to be the cause of phase variations near 0 or 180 degrees, therefore, they are given less weight in the computation []. These measures are used to investigate functional connectivity and brain network features.

$$wPLI = \frac{E[|I(Xij)|]}{E[|I(Xij)|]}$$

$$E[|I(Xij)|]$$

In the given equation, I is used to represent an imaginary component, Xij is cross-spectral density of two signals i and j and E is used to represent expected (mean) value.

[2.1.2.5] Network Analysis Graph Theory is a framework for analyzing and studying networks, which are made up of nodes (vertices) and edges (links) connecting them. Graph theory is used for the representation and analysis of complex links or connections in the brain networks. It comes under the category of whole network estimation which analyzes the entire brain network and examines the connection between all pairs of brain areas or nodes. Its goal is to characterize the overall brain network's global features and connectivity patterns. The network dynamics and network connectivity also come under whole network estimation. Network dynamics is the study of how the links or interactions within a network change over time. Researchers can learn more about how the brain's connection patterns change during different cognitive states or in response to external inputs by studying network dynamics. On the other hand, network connectivity is used by researchers to uncover and characterize functional connections or networks in the brain that are linked to particular cognitive functions or neurological diseases. Table 4 represents the studies that have employed time-frequency, functional connectivity and network theory-based analysis methods to investigate EEG signals. These studies are provided with number of participants, their findings, brain regions used and the criteria used for the classification of cognition. By reviewing earlier research findings (Table 4), the author can gain a better understanding of connectivity and network-related analysis approaches and gain significant knowledge about cognitive classification.

Table 4: Studies related to time-frequency, connectivity and graph theory-based analysis of EEG [48-52, 64, 67-72]

Abbreviations Used in Table 4: HC = Healthy controls; RTF=Relative time spent by EEG frequency bands ; ROR= Resting occipital rhythm; PDD = PD with demented patients; MCI=Mild-Cognitive Impairment; PLI = Phase lag index; MoCA = Montreal Cognitive Assessment; DSM=Diagnostic and statistical manual of brain dysfunction; MMSE = Mini-mental Score examination; HAMA = Hamilton anxiety rating scale ; HAMD = Hamilton depression rating scale; dwPLI = debiased weighted phase lag index				
Author	Participants	Regions used for analysis	Criteria used for classification of cognitive status	Main findings
Min Cai et al. (2021)	68 PD patients (MCI /Non-MCI)	The posterior middle frontal gyrus PMFG) region	• HAMA and HAMD scales used for anxiety and depression access	• The study revealed unusual functional interconnection patterns in theta frequency band, in the region of
Lior Molcho et al. (2022)	Senior participants with different cognition= 60, HC= 22	Not Specified	• Diagnosis is done on the basis of MMSE score	• The study found distinct activity patterns among young and older participants in various mental states.
Marta Kopanska et al. (2022)	This study involves 5 people with a generalized anxiety disorder(GAD).	Central Region	• EEG is recorded in the rest and open-eye state. Based on the international 10-20 system, the QEEG data were collected by monitoring all waves from central sites (C3, CZ, C4).	• The results indicate specific pattern amplitudes at brain areas C3 & C4. The QEEG research gives critical information on the particular brain wave patterns in individuals with GAD, allowing for an early and reliable diagnosis of issues.
Emad Arasteh et al. (2022)	PD=15, HC= 18	All regions	• Sparse discriminant analysis is used to discriminate between PD and HC	• This study finds strong connections between the learned latent features and clinical indices connected to motor symptoms in PD.
Ute Gschwandtner et al. (2023)	Non- demented PD=75, 72 HC	All Brain Regions	• MMSE Score, PD patients were followed over 3 years	• To analyze dynamic changes in EEG synchrony patterns, they developed a method termed "Time-Between- Phase-Crossing" (TBPC).

[2.1.2.6] Individual alpha frequency (IAF)

IAF is performed as a part of the FD. IAF analysis determines the peak (highest) frequency of the alpha frequency band. The alpha (8Hz-12Hz) rhythm is an oscillatory brain activity that is, usually located in the brain's posterior regions, especially when a person is comfortable and their eyes are closed. IAF is linked to cognitive abilities and it reflects certain properties of the brain's functioning. One of the studies discovered that the IAF is highly constant in healthy individuals up to the age of 80 years [53]. They discovered that the IAF cannot be modified by cognitive therapies. Cognitive treatments are activities or exercises that aim to improve cognitive abilities. Cognitive interventions refer to activities or tasks that are designed to improve cognitive abilities. Therefore, in the case of a healthy individual with good cognitive functioning, the IAF is expected to remain stable even after cognitive interventions are performed. IAF is one of the most used methods for examining subject-to-subject variation in EEG rhythms. There are many different approaches in the literature for determining IAF, including the well-known popular peak- frequency (PF) technique, the transition-frequency (TF) technique and the extended-band (EB) technique. Although based on qualitative criteria, it could be challenging and time-consuming to repeat since it requires a visual evaluation of each individual EEG spectrum. In order to get around this issue, earlier research proposes a novel channel reactivity- based (CRB) technique for determining IAF [54]. Because of its emphasis on quantitative indications and criteria, CRB focuses on task-driven alpha response patterns rather than the existence of EEG peaks. Table 5 summarizes the investigations and findings based on IAF analysis methods of EEG signals.

Table 5: Studies based on the IAF analysis method of EEG [20-22]

Abbreviations are: HC = Healthy controls; IAF = Individual Alpha Frequency; MCI=moderate Cognitive Impaired patient; DLB = Dementia with Lewy bodies; ADD = AD patients with dementia; PDD = PD with demented patient;				
Author	Participants	Regions used for analysis	Criteria used for classification of cognitive status	Main findings

Babiloni et al. (2017b)	HC = 75, AD-MCI = 75, PD-MCI = 75	Frontal, Central, Parietal, Temporal, Occipital	• AD-MCI was diagnosed with a resting-state EEG with MMSE score of 24, • A logical memory tests Score • A DSM criterion is used for PD-MCI	• Author examined IAF and found differences in the groups of ADD, PDD, DLB and healthy controls. • The order is (HC > AD > PD). The PDD and DLB had the lowest IAF values, while the HC group had the greatest IAF
Olde Dubbelink et al. (2013)	HC = 14, PD = 43	All regions	• Eyes-closed resting state condition • Cambridge Cognitive Examination (CCE) was used to evaluate global cognitive function and detect the presence of dementia.	• This study lasted four years and was a longitudinal study • The major peak frequency of IAF was greater in HC as compared to PD. (HC > PD) • HC IAF remained largely stable between baseline and follow-up, but PD group experienced, on average, lower IAF than baseline

Based on the prior research, spectral analysis may reveal the EEG signal's frequency content, but it does not give the information that how these frequency components change over time. Similarly, IAF analysis provides information about the peak alpha frequency in an individual, but it does not capture the changes in alpha activity over time. Therefore, time & frequency analysis is an important tool in EEG analysis because it offers a more comprehensive understanding of the rapid alterations in neurological activity that occur throughout time. Another key component of EEG analysis is investigating brain connection patterns. Graph theory-based network analysis and functional connectivity measurements may help in determining the connections between various areas of the brain.

These techniques enable researchers to better understand the complex relationships and interactions that exist between different brain regions or electrodes. Therefore, researchers can gain a deeper understanding of neural mechanisms and identify potential biomarkers for the examination and treatment of PD by selecting a suitable analysis method aligned with their research objective.

[3] EXPERIMENTAL ANALYSIS OF ML MODELS FOR CLASSIFICATION OF PD-D, PD-MCI, PD-NC & HC

Various ML, DL and TL approaches can be used to predict and classify PDD, PD-NC, HC and PD-MCI categories. There are numerous strategies available in the literature [55-59]. In the present work, some ML methods have been used to classify PD-NC, HC, PD-MCI, & PDD. The dataset has been obtained from the North Shore Institute of Health Science, Chicago. The purpose of this experiment is to find the best classification approach for accurately classifying cognition using EEG signals. The results of the techniques have been presented and tabulated for comparison.

[3.1] Objectives of this study are.

- Comparative analysis of ML algorithms for predicting different stages of cognition
- Spectral & temporal analysis for feature extraction of the Parkinson's disease stage prediction i.e. HC, PD-NC (high cognition), PD-MCI (intermediate cognition), PD-D (low cognition)

Various steps that lead to the classification are followed in next section (Fig. 4).

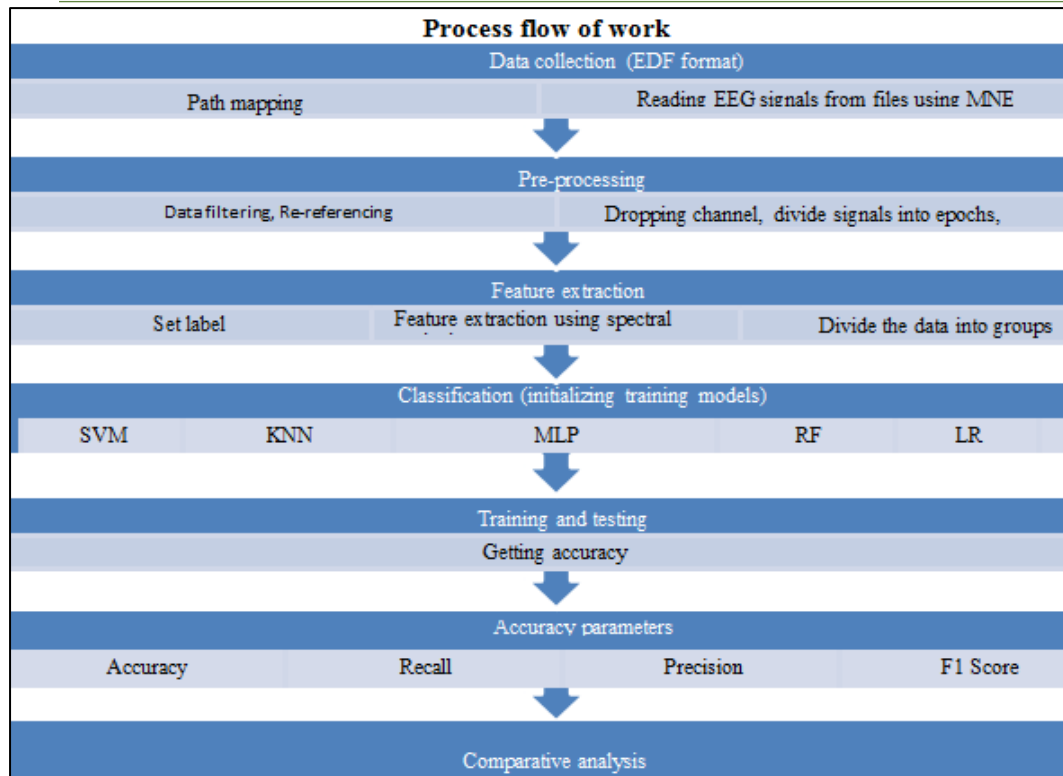


Fig. 4: Main steps of work process flow

Various steps that lead to the classification are followed in the next section. In Fig. 4, to get noise-free and artefact free information from EEG signals, the preliminary stage is pre-processing. Pre-processing of these signals includes data filtering, re-referencing, removing bad channels, dividing the signals into epochs and using ICA to get rid of artifacts. After pre-processing of signals, labels are assigned to the various categories (PDD, HC, PD-MCI, PD-NC) based on MoCA and UPDRS score. The next phase is feature extraction, which extracts important characteristics from the pre-processed EEG signals using spectral and statistical evaluation approaches. These obtained features are then used to train classifiers such as SVM, KNN, MLP, RF and LR. The trained classifiers are subsequently tested and the results are presented in a tabular format. Various evaluation metrics (accuracy, F1-score, precision, recall and AUC score) are used to compare the results of different classifiers.

[3.2] Dataset used

The North Shore institute (Chicago) makes the dataset available upon request through the appropriate channels [60]. For the experiment, a total of 40 persons (healthy as well as PD patients) of the same sex & age group are used. Standard clinical evaluation of PD patients includes obtaining scores on the UPDRS-III scale, disease-stage scale Hoehn and Yahr (H&Y) & Montreal Cognitive Assessment i.e. MoCA recommended scale for measuring cognitive changes in PD. The UPDRS is a standardized scale used to examine and monitor the severity of PD symptoms. MoCA is a quick way to gauge a patient's mental health. The MoCA measures a variety of mental abilities, such as focus, spatial awareness, memory, language and decision-making. A score of 30 indicates excellent cognitive function; the test takes around 10 minutes to perform on average. UPDRS and MoCA are both scoring systems used in medical assessments, they serve different purposes. The UPDRS is used specifically for assessing Parkinson's disease symptoms, while the MoCA monitors cognitive skills in patients.

In this paper, the dataset is divided into four categories based on the MoCA scores, a PD-NC MoCA score is 26 or above; PD-MCI is defined as a score between 19 and 25 and PDD is defined as a score of 18 or below [61-62].

[3.3] Methodology

To highlight the efficiency of different ML methods, five different supervised classifiers are considered: Support Vector Machine (SVM), Logistic Regression (LR) & K-Nearest Neighbor (KNN), Multi-Layer Perceptron (MLP) and Random Forest (RF) classifiers. Finally, a comparative analysis of these techniques is done.

[3.3.1] Data pre-processing

Data preprocessing is the most necessary step of EEG analysis which makes the data usable. The dataset includes non- EEG parameters and there are total of 51 channels, but only 20 channels have been employed for the study as all other channels are not mapped according to the “10-20” International Montage System. A band-pass filter is used to process data with the frequency ranges from 0.5 Hz and 45 Hz. Frequencies higher than 45 Hz with line noise & artifacts and low frequencies below 0.5 Hz, are also filtered. The signal is then normalized via re-referencing. ICA is applied to remove ECG and EOG artifacts. ECG artifacts are heart-related and EOG artifacts are related to eye blinking, eye movement, etc. To enhance the categorization and interpretation of the EEG data, signal segmentation is carried out. The whole 3 minutes' time series of the EEG signal is split up into units known as epochs. Rest of the steps after preprocessing are feature extraction and classification, which are explored in the proceeding subsections.

[3.3.2] Feature extraction

Feature extraction is done using spectral analysis and statistical methods, after combining all these features, the classifiers are trained for the classification. Fig. 5 illustrates the feature extraction procedure for one EEG sample using spectral analysis. The EEG data is separated into segments of 5 seconds after preprocessing. After signal segmentation, the Discrete Fourier Transform (DFT) is applied to each unit individually to transform TD signal to FD signal. It provides the segmented EEG data's amplitude and phase information at various frequencies.

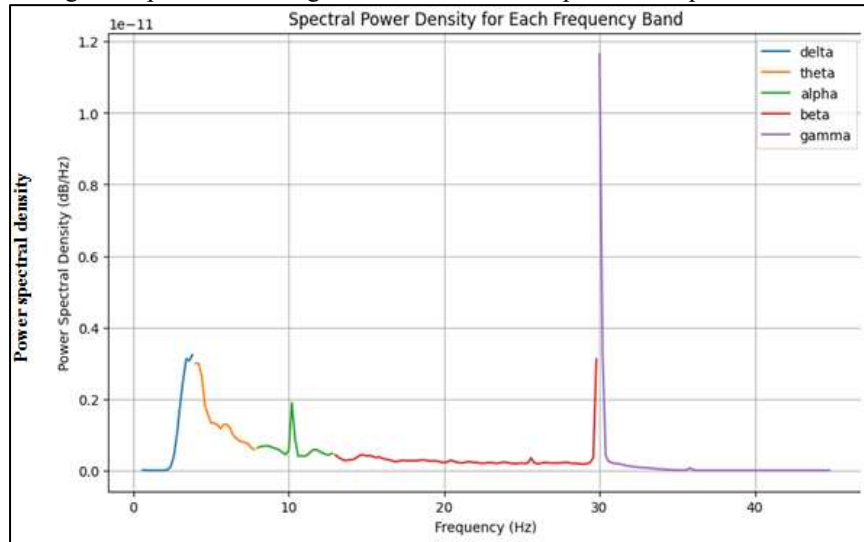


Fig. 5: Steps for feature extraction using spectral analysis

(a) Spectral power

Spectral power density is computed with the Welch technique to calculate the spectral properties. From the total power of the signal, it is possible to determine the power of each frequency band using the power spectral density. To analyze the strength distribution over the neurological frequency spectrum, the EEG data is separated into five frequency bands. For EEG data segmentation, the welch method relies on Discrete Fourier Transform (DFT) to perform periodogram (TD signal to FD) calculations as displayed in Fig. 5. The windowed data is transformed using a DFT to provide the overlapping segment periodogram. The power measure is then calculated by squaring the data and averaging the periodograms.

(b) Statistical methods

Some temporal characteristics include the mean, peak-to-peak (PTP), variance, maximum, minimum, standard deviation (SD), index of maximum value, index of minimum value & absolute difference of signal, as well as skewness & kurtosis, are extracted using statistical methods.

[3.3.3] Classification

- **Support Vector Machine (SVM):** The linear and nonlinear issues can both be handled by the SVM algorithm. The technique's main goal is to find the line, or hyperplane, that best divides the categories. The largest margin between the values on either side of the decision line is the main focus of SVMs. The primary objective of the technique is to locate the line (hyperplane) that optimally separates the categories.
- **K-Nearest Neighbor (KNN):** The KNN is an example of an irregular method used in ML. New data items are classified in accordance with their similarities to records in the training set. The similarity is often defined in terms of the Euclidean distance or another commonly used metric. When making predictions, the approach finds the most similar existing record and gives that record's classification to the new record. The most popular class is chosen from among a group of 3, 5 or 9 neighbors. KNN classifier's default K value is 5.
- **Multi-Layer Perceptron (MLP):** MLP classifier uses supervised learning, it needs a training dataset in order to learn. The training dataset includes output labels and input data (x) (y). Using the input data, the MLP classifier learns how to translate the data to the output labels. The MLP classifier may be trained to convert data into labels and then used to make predictions on fresh data. This classifier takes in data (x) and outputs a label prediction (y). As a robust ML strategy, it may be put to use in both regression and classification. It is often used in applications like automatic speech and image recognition.
- **Random Forest Classifier (RF):** The supervised learning framework from which RF emerges. It can be applied to ML problems involving regression and classification. Its foundation is the idea of ensemble learning, which integrates a number of classifiers to tackle intricate issues and boost model effectiveness.
- **Logistic Regression (LR):** Making predictions regarding a dependent or target variable is done using this supervised learning technique. Correlations between the variable that is dependent and several independent variables are found using LR. For binary categorization, LR models are frequently utilized.

From the literature, ML techniques or hybrid ML techniques have been used due to the short size of the dataset in this research field. However, to get better and more accurate results, the use of DL methods such as CNN and RNN networks can be investigated further. These advanced methods can help to understand the cognitive patterns in PD patients more effectively.

- **Deep Learning (DL):** A subfield of ML called DL focuses on training multi-layered artificial neural networks. These neural networks can perform difficult tasks like natural language processing, image and audio recognition and more because they can learn data with the hierarchy structures. These are a few categories of DL.
- **Convolutional Neural Networks (CNN):** CNNs are a subset of DL architecture. These are used for processing grid-like input, including images or 2D signals. CNNs do well in finding and assessing spatial trends in data. They employ specialized layers, such as convolutional, pooling and fully linked layers.
- **Recurrent Neural Networks (RNN):** RNNs are a form of DL architecture that is specifically developed for processing sequential input with temporal relationships. These are the best structures for dealing with time series data. Unlike feedforward networks, RNNs feature connections that allow information to persist and be modified over time. RNNs are suitable for applications such as recognition of speech & processing of natural languages, because they process incoming data sequentially and retain information from previous inputs. Because RNNs have internal memory states, they can access and recollect past data to influence their current forecasts. A couple of RNN modifications that enhance the ability to detect long-term dependencies are the Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM).
- **Transfer Learning:** To overcome the problem of having a limited amount of data, an approach called Transfer Learning (TL) can be explored. This approach makes use of models that have already been trained on large datasets from related areas, such as VGG16, VGG19, Inception V3, Resnet50, etc.
- Simulation of work has been performed using Python 3.9 on Google Colab. EEG data is in EDF format, EDF stands for "European Data Format," a file format that stores information in 16 bits. MNE library is used to access and edit EDF files. MNE library includes a high-level interface for easy access to read and write EDF files. EDF browser is used to visualize all the EEG signals.

[4] RESULTS AND DISCUSSION

[4.1] Statistical Analysis

Table 6 displays the baseline clinical characteristics and demographics of the four study groups. All the groups (HC, PD-NC, PD-MCI, PDD) are notably different from each other in terms of motor score (UPDRS), cognitive score (MoCA) and disease duration. These groups are compared statistically on the basis of the average values with their standard deviations. The means of several groups are compared using the ANOVA test to see if there are any statistically significant differences between them. The ANOVA test is used to compare the means of multiple groups to determine if the differences between them are significant. Except for age, all the groups showed significant differences. In the ANOVA test, F and P values are obtained, where 0.05 ($P < 0.05$) was chosen as the statistical significance criterion. The p-values for UPDRS score, PD duration and MoCA score were 0.001, which is significantly lower than the predefined significance criterion of 0.05. This indicates that there are substantial differences between the categories in terms of these measures. However, the results for age were not significant, suggesting that there are no differences in age across the categories.

Abbreviations used: N/A=Not applicable; UPDRS III=Unified PD Rating Score (Part III to check the level of motor symptoms); N=Number of candidates; MoCA=Montreal Cognitive Assessment (level of cognition), $p < 0.05$

Table 6: Mean (SD) of baseline clinical and demographic characteristics of categories

	HC, N=18	PD-NC, N=6	PD-MCI, N=7	PDD, N=7	F-value	P-value
Age(years)	72(8.5)	71(7.0)	73(5.5)	73(10.2)	0.166	0.918
Sex(F/M)	7/11	4/2	2/5	2/5	-	-
UPDRS III score	N/A	15(9.6)	15(7.1)	23(7.1)	25.690	<0.001
MoCA Score	30(0)	30(1.1)	22(2.3)	12(4.7)	28.313	<0.001
Duration of PD	N/A	4(1.3)	5(4.5)	5(2.5)	13.892	<0.001

The results from using traditional ML and ensemble approaches with a different set of features are analyzed and discussed in this subsection. After model-specific hyper-parameter tuning to enhance model performance, all of the results are mentioned. F1 score and AUC-ROC have been selected as the main techniques of evaluation since they together offer a reliable method for determining predictive performance. The relative strengths in every frequency range for all the four groups are used to conduct the classification. When compared to the other ML techniques (SVM-78%, KNN-81%, MLP-75%, Logistic Regression-72%) in this experiment, the RF approach yields the highest accuracy (91%) to identify four categories of cognition (PD-NC, PD-MCI, PDD, HC) in PD.

The Logistic Regression classifier has a 72% overall accuracy for classifying various categories of cognition and its hyperparameters are "C" = 0.0001, 0.001, 0.01, 0.1, 1 and "penalty" = l2. Table 7 compares the accuracy, precision and F1 score for Parkinson's disease (NC, MCI, PDD) patients with healthy controls in the same sex and age range. With 93.14% accuracy and a precision rate of 0.76, LR has the highest classification accuracy for the PD-NC group but the lowest classification accuracy for the healthy control groups. The average curve area (AUC) for LR according to the ROC curve (Fig. 6) is 0.90.

The AUC score ranges from 0 to 1, with 1 indicating the best possible classifier and 0.5 indicating the need for improvement.

The model has a 90% chance of properly ranking a positive instance higher than a negative instance, according to

an AUC score of 0.90. However, compared to other classifiers, this one has the lowest overall accuracy. Due to the non-linear character of EEG signals, LR, which is a linear model, may not properly capture this non-linear nature. As a result, its accuracy is lower than that of more sophisticated non-linear models.

Table 7: Logistic Regression

Class	Accuracy	Precision	Recall	F1 Score
Controls	75.52%	0.82	0.73	0.77
PD-NC	93.14%	0.76	0.70	0.73
PD-MCI	88.36%	0.59	0.81	0.68
PDD	86.87%	0.55	0.61	0.58

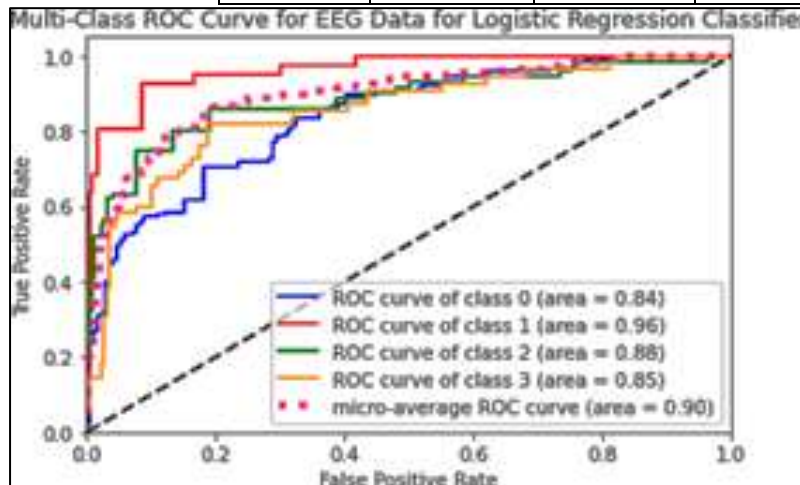


Fig. 6: AUC plot for Logistic Regression with spectral and statistical features

The overall accuracy for the classification of different categories of cognition using the KNN classifier is **81%** with the relative power across five frequency bands along with the statistical features. The hyperparameters used for the KNN classifier are 'n_neighbors'=7,5,3, 'weights'=uniform, distance and 'p'=1,2. Table 8 shows accuracy, precision

& F1 score for healthy controls and PD patients (NC, MCI, PDD), with same-sex and age groups for KNN. This classifier gives the highest accuracy to classify the PD-MCI group with 94.03% accuracy with a precision rate of 0.82, whereas it gives the lowest accuracy in predicting Healthy Control groups. ROC curve (Fig. 7) shows the average curve area for KNN is 0.94, which means that the KNN model performs very well in distinguishing between the classes. KNN is a non-parametric classifier, which performs well when the number of features is relatively small and it can handle non-linear relationships between features and target variables more effectively. Therefore, this classifier performs better than LR, SVM and MLP classifiers.

Table 8: KNN

Class	Accuracy	Precision	Recall	F1 Score
Controls	84.18%	0.85	0.84	0.84
PD-NC	91.34%	0.83	0.61	0.70
PD-MCI	94.03%	0.82	0.89	0.85
PDD	91.64%	0.65	0.80	0.72

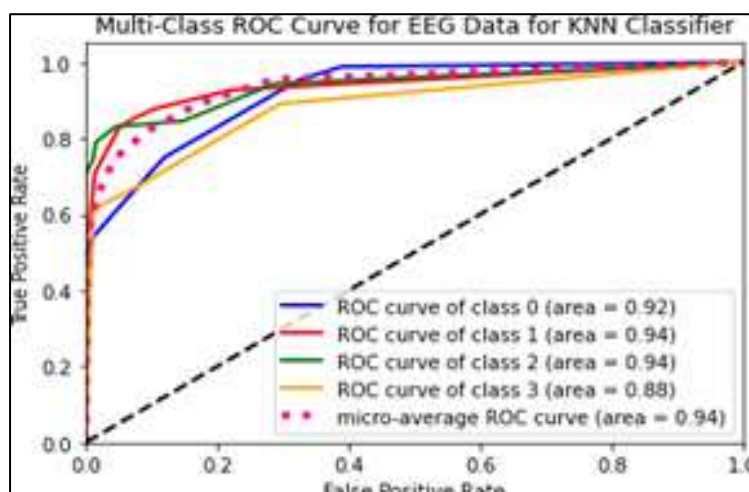


Fig. 7: AUC plot for KNN Classifier with spectral and statistical features

With spectral and statistical information, the MLP classifier's total accuracy is 75%. The hyperparameters for the MLP classifier are adaptive learning rate, 1000 maximum iterations and 100 hidden layer sizes. Table 9 shows the

accuracy, precision & F1 score for matched age and sex groups made up of PD patients (NC, MCI, PDD) and healthy participants. MLP predicts the PD-NC group with the highest accuracy (95.52% accuracy and a precision rate of 0.88), but predicts the Healthy Control groups with the lowest accuracy. The ROC curve (Fig. 8) shows that the average curved region for MLP is 0.92. In classification tasks, MLP perform better than LR classifiers because they can capture the non-linear relationships between the input characteristics and the desired output, while LR classifiers retain a linear relationship.

Table 9: MLP

Class	Accuracy	Precision	Recall	F1 Score
Controls	87.16%	0.87	0.87	0.82
PD-NC	95.52%	0.88	0.78	0.76
PD-MCI	90.45%	0.75	0.79	0.66
PDD	88.06%	0.64	0.64	0.60

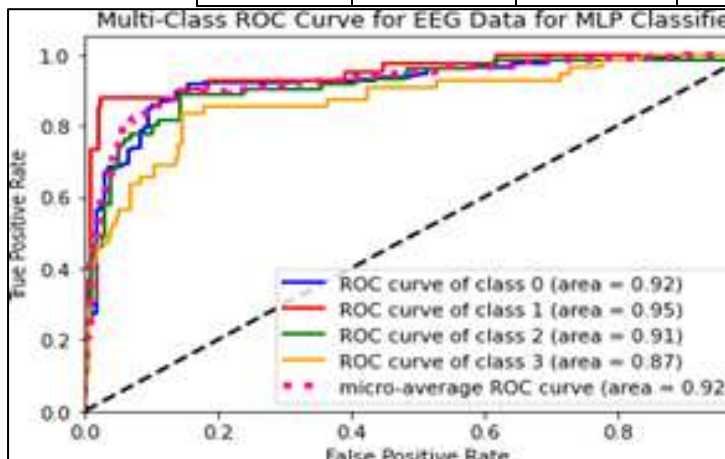


Fig. 8: AUC plot for MLP Classifier with spectral and statistical features

The highest accuracy (91%) is achieved by RF. Table 10 shows accuracy, F1 score, precision for sex and age-matched, healthy individuals and PD patients (NC, MCI, PDD) using Random Forest. The hyperparameters used for the RF classifier are 'n_estimators'= 50, 'min_samples_leaf'= [3, 4] and 'max_depth'= 4, 5. It gives the highest accuracy to classify PD-NC group with 99.1%, whereas HC group with 91.04%, PD-MCI with 95.22% and PDD with 96.72% accuracy and the highest precision rate is 0.99. ROC curve (Fig. 9) shows the average curve area for Random Forest is 0.99. An AUC score of 0.99 suggests that the model has a 99% chance of correctly choosing a positive instance higher than choosing a negative instance. This classifier outperforms all other classifiers to classify the cognitive status of patients because it effectively handles the complex and nonlinear patterns present in the EEG data related to the cognitive processes, due to which RF gives the best performance among all other models.

Table 10: Random Forest

Class	Accuracy	Precision	Recall	F1 Score
Controls	91.04%	0.99	0.85	0.92
PD-NC	99.1%	0.93	1.0	0.96
PD-MCI	95.22%	0.77	1.0	0.87
PDD	96.72%	0.82	0.98	0.89

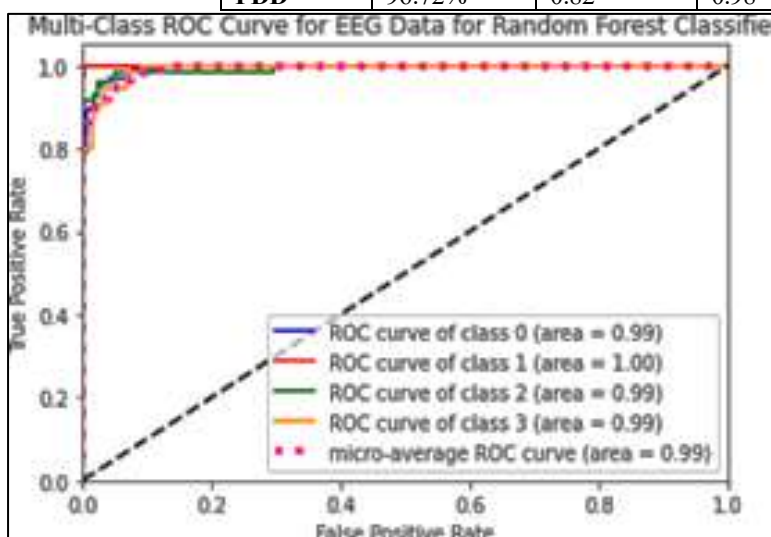


Fig. 9: AUC plot for Random Forest Classifier with spectral and statistical features

The overall accuracy of SVM classifier is 77% with the spectral and statistical features. The hyperparameters used

for the SVM classifier are 'C'= 1,10 and 'kernel= linear'. Table 11 shows accuracy, precision and F1 score for patients with PD (NC, MCI, PDD) and healthy controls, in the same sex and age group using SVM. It gives the highest accuracy to classify PD-NC group with 94.63% accuracy and a precision rate of 0.85, whereas it gives the lowest accuracy in predicting Healthy Control groups. ROC curve (Fig. 10) shows the average curve area for SVM is 0.87, which means that classifier needs to improve its performance. This classifier performs better than LR and MLP classifiers because SVM's ability to capture and model nonlinear relationships can lead to better performance compared to LR and MLP classifiers, which do not capture such complexities so effectively.

Table 11: SVM

Class	Accuracy	Precision	Recall	F1 Score
Controls	83.88%	0.83	0.84	0.84
PD-NC	94.63%	0.85	0.74	0.80
PD-MCI	88.66%	0.76	0.72	0.74
PDD	87.46%	0.55	0.64	0.59

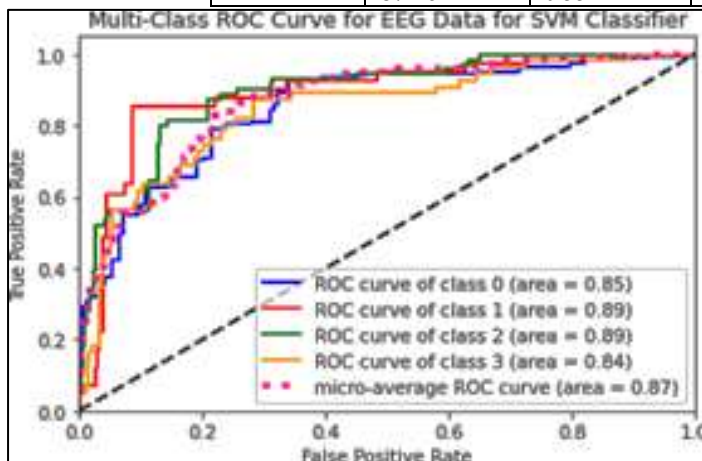


Fig. 10: AUC plot for SVM Classifier with spectral and statistical features

Table 12 shows the Accuracy chart for all categories with different classifiers. In Fig. 11, Random Forest gives the best accuracies for all categories. Among all the groups, PD-NC are classified with the highest accuracy by all classifiers. All the accuracies are above 70% for all classifiers. Healthy Control groups are classified with the lowest accuracy among all other groups.

Table 12: Accuracy comparison chart

	SVM	KNN	MLP	RF	LR
Controls	83.88%	84.18%	87.16%	91.04%	75.52%
PD-NC	94.63%	91.34%	95.52%	99.1%	93.14%
PD-MCI	88.66%	94.03%	90.45%	95.22%	88.36%
PDD	87.46%	91.64%	88.06%	96.72%	86.87%

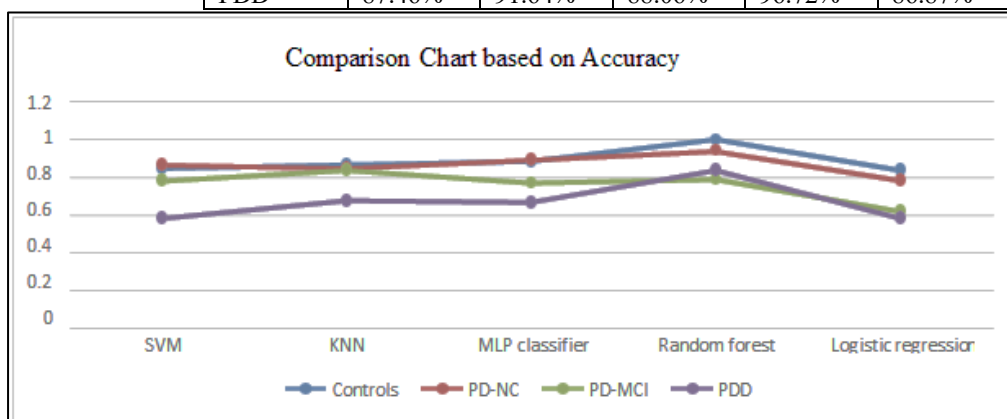


Fig. 11: Comparison chart on the basis of Accuracy

Table 13 shows the F1-Score Comparison Chart for all categories with different classifiers. In Fig. 12, the Random Forest gives the highest F1 score for PD-NC group, whereas KNN classifier shows the lowest F1-score for PD-NC group. KNN gives the best F1-score as 1 for PDD group, indicating the perfect precision and recall. Overall Random Forest shows the best score (close to 1) for all categories.

Table 13: F1-Score comparison chart

	SVM	KNN	MLP	RF	LR
Controls	0.84	0.84	0.82	0.92	0.77
PD-NC	0.80	0.70	0.76	0.96	0.73
PD-MCI	0.74	0.85	0.66	0.87	0.68

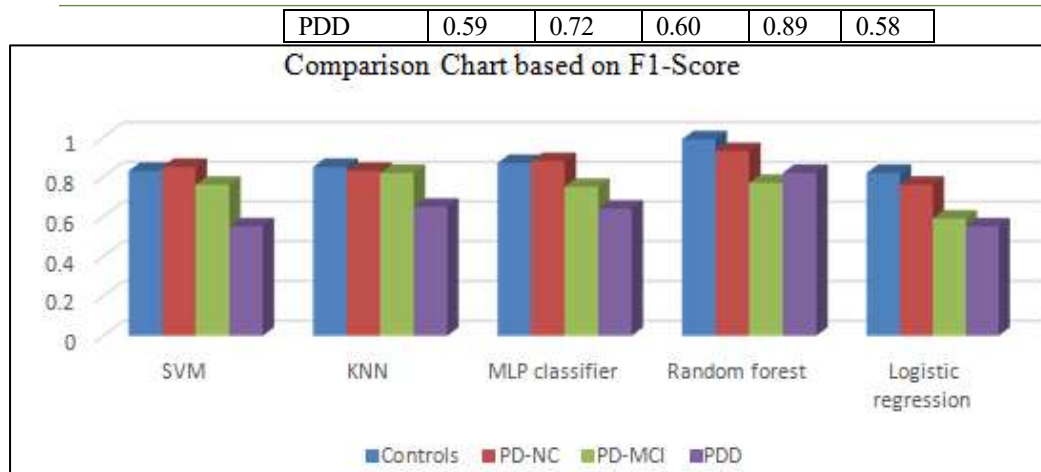


Fig. 12: Comparison chart on the basis of F1-Score

Table 14: Precision comparison chart

	SVM	KNN	MLP	RF	LR
Controls	0.83	0.85	0.87	0.99	0.82
PD-NC	0.85	0.83	0.88	0.93	0.76
PD-MCI	0.76	0.82	0.75	0.77	0.59
PDD	0.55	0.65	0.64	0.82	0.55

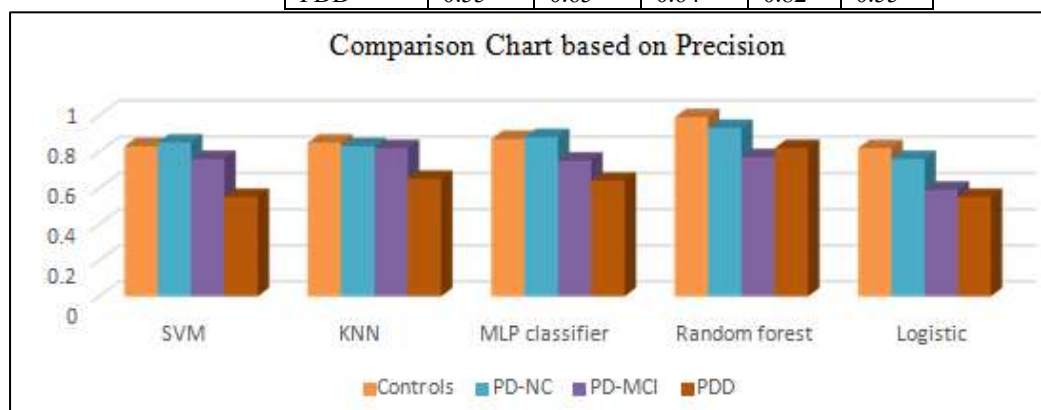


Fig. 13: Comparison chart on the basis of Precision

On the basis of these comparison charts, Random Forest outperforms the other classifiers in Accuracy, F1-score and Precision (Fig. 13 & Table 14). Random Forest effectively captures the complex and nonlinear patterns present in EEG data related to cognitive processes, enabling accurate classification and interpretation of cognitive states. RF has the ability to handle nonlinear relationships and it is robust to noise & outliers in EEG signals.

[4.2] Prisma Model

In the flow diagram of Prisma model (Fig. 14), the three primary steps of the search strategies are identification, screening and eligibility. The identification stage is the first step in gathering all possibly relevant studies which address the research problem. A thorough search is carried out across many databases, employing keywords, subject headings and other search parameters. The primary purpose at this stage is to find as many studies as possible that could be considered in the review, without regard for whether they fit the criteria. Following the identification stage, numerous research are collected. The screening step involves eliminating studies written in non-English languages and extraneous book chapters. Titles and abstracts are quickly evaluated to exclude irrelevant research, duplicate entries and studies which fell beyond the scope of the review. Only studies that passed the initial screening, proceeded to the next round. At the eligibility step, all remaining studies' full texts are carefully examined to make sure they satisfied the comprehensive inclusion requirements for the review. Compared to the original screening, this step offered a more thorough examination. The final synthesis of the review then contained the studies that are found eligible after going through this procedure. In the included stage, 59 articles are chosen for review based on cognitive classifications in PD. These studies are organized as follows: 21 research compare MCI with non-MCI, while 11 studies distinguish between HC, MCI and non-MCI. Eight research investigate MCI and PDD. Three studies examine HC, PD-NC, MCI and PDD. Eleven research compare PD-NC, HC and MCI, while five studies categorize PD-NC, MCI and non-MCI [4,6-7,10-16,19-25,36-40,43-50,53-55,57,61,63-75,77-87].

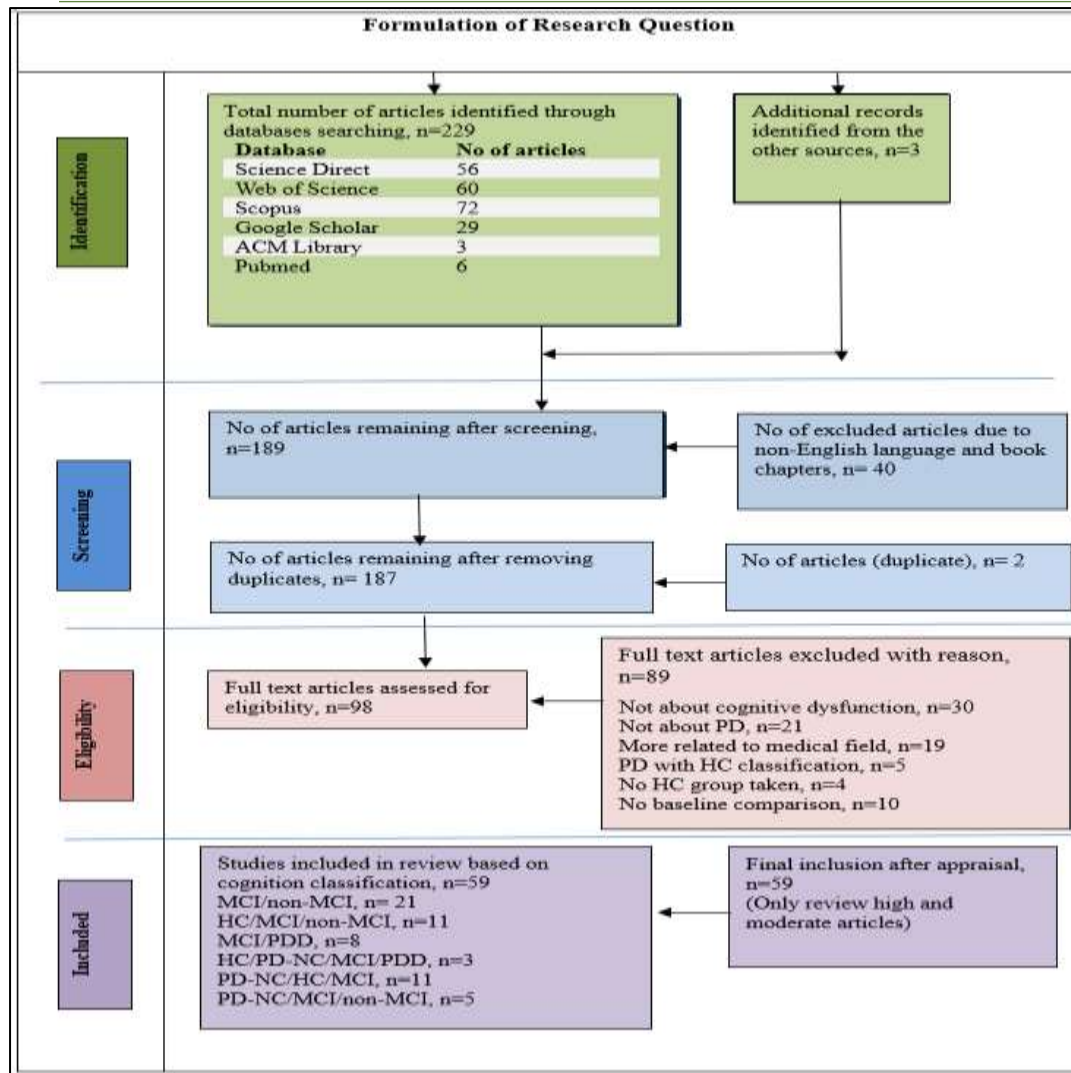


Fig. 14: The flow diagram of PRISMA Model

The graph (Fig. 15) illustrates a comparison of studies conducted from 2016 onward, highlighting the accuracy achieved in diagnosing cognitive decline in PD using EEG data. The studies use various ML approaches and the accuracy ranges from 80% to 92%. Notably, [25] achieved the best accuracy of 92% by combining the Tsfresh lib+ML pipeline with Occipital peak frequency. However, that high accuracy is limited to identifying only two cognitive groups. The general pattern shows that improvements will occur over time, particularly with the combination strategies, which consistently outperform single methods. Current approaches still fall short of providing complete and highly accurate classifications for a full spectrum of cognitive disorders in PD. Therefore, to more accurately classify the entire spectrum of cognitive impairment, more advanced and novel approaches must be developed and integrated.

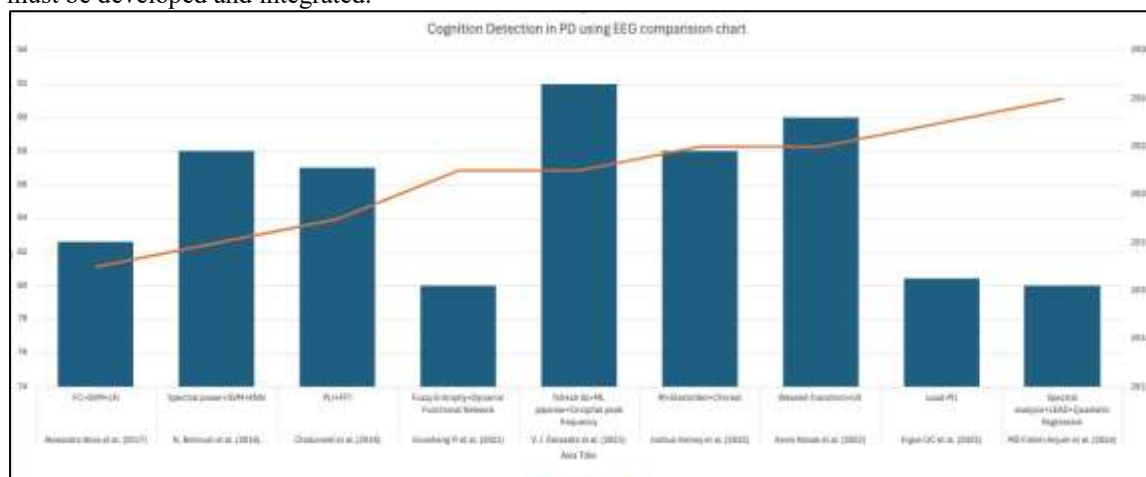
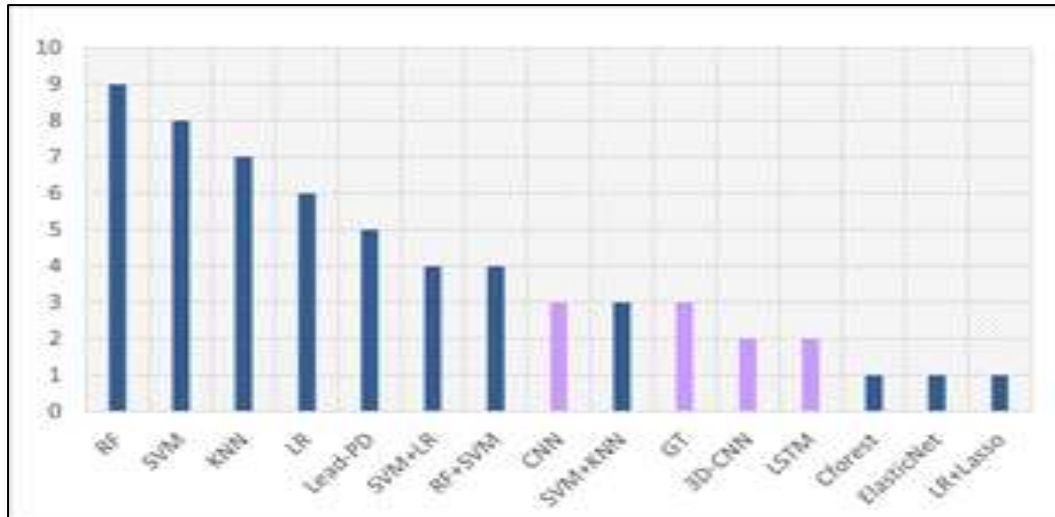
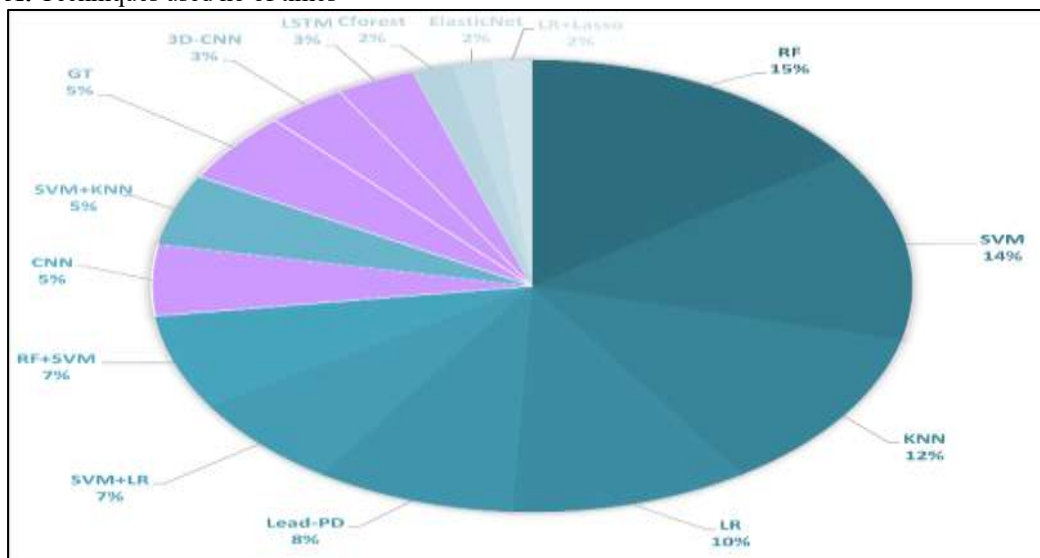


Fig. 15: Comparison chart on the basis of accuracy for cognition detection in PD using EEG with techniques used



A: Techniques used no of times



B: Types of Models used

Fig. 16: A) Frequency of ML models in EEG-based cognition detection studies for PD, B) Model type distribution in PD cognition detection: comparative usage rates of non-neural and neural network techniques in 59 studies.

In Fig 16. Figure A presents a bar chart illustrating the frequency of various models used in the analyzed studies to diagnose cognitive impairments in PD using EEG data. The analysis shows that RF and SVM were the most commonly used models, accounting for 14% of the 59 publications reviewed [21, 25,36-40,48, 78]. These were followed by KNN, which was utilized in 12% of research, LR in 10% and the LEAD-PD model in 8% of the studies [63-69,71-72,78-80,84]. Figure B depicts a pie chart categorizing the models used in 59 selected studies as non-neural machine learning (ML) techniques or neural network-based approaches. Notably, 11.5% of the models used in this research were neural network-based, with DL approaches, including CNNs, being commonly used. However, 88.5% of the models used were non-neural ML approaches.

[4.3] Discussion and Future Directions

From 2006 to 2024, multiple studies were conducted to diagnose brain dysfunction in PD utilizing EEG and ML techniques. These investigations used EEG data to generate accurate cognitive biomarkers for PD patients, allowing for more effective monitoring of cognitive function. However, the classification of cognitive impairment has primarily relied on machine learning techniques due to the small size of medical datasets, which poses a significant challenge. DL approaches have been utilized extensively in numerous EEG-based areas, including emotion recognition, schizophrenia detection, epilepsy prediction and brain-computer interface (BCI) applications. However, its use in detecting cognitive decline in PD is relatively unexplored. While machine learning methodologies are frequently employed, only a few studies have explored deep learning for this purpose, highlighting a significant research gap [81-85]. One study used a 1D CNN-LSTM architecture to extract salient characteristics from EEG signals, resulting in excellent accuracy when categorizing PD patients with cognitive deterioration [81,82]. Another study employed 3D- CNNs on sleep EEG data and achieved promising results in identifying PD patients with MCI and NC [83].

When working with the EEG data, the challenges associated with EEG are unwanted noise, artifacts from many sources such as muscle movements, eye blinks, electrical interference and non-stationary (properties change over time) properties, which can impact data quality and analysis approaches in EEG signals. During EEG acquisition,

the signals may change from one individual to another due to variations in skull thickness, structure of brain & conductivity. This condition is called data variability which is the most common problem with EEG data [41]. Furthermore, EEG electrodes record signals from many brain locations, which makes it difficult to identify specific signal sources. Improved source localization approaches help researchers and clinicians to provide more precise details of the areas of brain that are involved in specific cognitive processes. This method can better estimate the locations of neural activity within the brain. Advanced methods are frequently used in these techniques to improve the accuracy of source localization. EEG analysis includes laborious preprocessing and manual feature selection techniques which requires a lot of expertise. It requires an excessive amount of time and work to manually choose and compute the features from EEG data, hence an automatic approach is required to carry out each step for all EEG signals. Therefore, For the future, ML and DL technologies may be utilized to analyze EEG signals in place of traditional approaches to automate the EEG analysis process. These methods can be used for numerous purposes like feature extraction, identification, pattern recognition and data denoising, etc. In this paper, the purpose of this experiment and comparison of various classification techniques is to present the best method among various techniques available in the literature. The aim of this experiment is to find an effective classifier that could deal with the complexity and non-linearity of EEG signals. In this experiment, feature extraction is done by combining spectral characteristics and statistical features. Spectrum indicators of cognitive impairment may be seen in the delta (0.5-4.0Hz), theta (4.0-8.0 Hz), alpha (8.0-10.0 Hz) and beta (13.0-30.0 Hz) frequency range in PD patients. The classifiers such as LR, SVM, MLP, KNN and RF are thoroughly analyzed. The findings show that the RF classifier surpassed the rest of the classifiers, attaining an impressive 91% accuracy. This study's findings not only increase present diagnostic practices but also influence future research efforts, helping both clinicians and scientists in the development of more accurate cognition classification techniques. One of the major drawbacks of EEG-based cognition classification is the small dataset. Two strategies—data augmentation (DA) and transfer learning (TL)—have drawn a lot of attention in the literature as potential solutions to this problem. The process of creating additional training data from the old data after using various DA approaches is known as data augmentation. Conversely, transfer learning fine-tunes pre-trained models which have already been trained on relevant EEG datasets for the particular cognition categorization task. These strategies may provide potential solutions to the problem of data scarcity in EEG-based cognitive classification. Future studies could use advanced techniques, such as directional functional connectivity, to examine the neural networks involved in cognitive processes in PD. In the frequency domain, bi-spectral analysis is also an effective tool for researching cognition in PD. This method can be used by researchers if they want to deeply investigate nonlinear interactions and complicated dynamics in EEG data, particularly when looking into cognitive functions and neurological disorders. Bi-spectral analysis has been employed in EEG for BCI and emotion identification tasks [39, 42], but its potential for classifying cognition in PD has not yet been investigated. Additionally, this paper's experimental results demonstrate that ML algorithms have the capability to classify cognitive aspects in PD with a high degree of accuracy. However, it is also important to investigate deep learning models and neural networks like GNN (Graph Neural Networks), which have shown good accuracy and performance in BCI, seizure detection and emotion classification tasks using EEG [73-75]. GNNs are a type of artificial neural network that is designed to deal with networks or graphs. GNNs have shown potential in tasks such as functional connectivity analysis, brain network analysis and EEG data classification, making them a useful tool in EEG-based research. For future research, source-level EEG analysis may provide more accurate and detailed information than sensor-level data, providing deeper insights into cognitive processes. Furthermore, rather than assessing all brain regions, concentrating on specific regions of interest may produce more targeted and useful results in the context of cognition detection in PD. Despite this, there has been a very modest number of connection studies that have specifically addressed the cognitive states of PD patients [63-66, 70-72, 76-77]. This indicates that there is still much to be learned regarding the brain processes that underlying cognitive dysfunction in PD and there is a need for more research in this area.

5. CONCLUSION

In this paper, a comprehensive review of various pre-processing, analysis and classification techniques is given. Early identification of several illnesses, such as cognitive impairment in PD and AD, is a challenging and complex task. Therefore, in this paper the objective is accomplished using traditional ML methods such as MLP, RF, KNN, LR & SVM, each of which has an F1 score ranging from 60 to 90%, an accuracy of between 72 and 91% and precision ranges from 50 to 100%. The best outcome is obtained during Random Forest classification to classify the cognitive status of four categories (HC, PD-NC, PD-MCI, PD-D). Random Forest, KNN, SVM, MLP and Logistic Regression are the classification algorithms in descending order of maximum accuracy attained. Random Forest effectively captures the complex and nonlinear patterns present in EEG data related to cognitive processes. As a result, RF outperforms all other classifiers in the present case. To acquire the best outcomes from the Classifiers, hyper-parameter optimization is also used. Therefore, this study's practical contribution to the field of neuroscience is to use spectral analysis as a feature extraction method and find the optimum classification algorithm to categorize the cognitive profile of PD patients using EEG signals. Furthermore, A total of 59 previous researches are analyzed and it is discovered that non-neural techniques were more frequently used, often in combination with other feature extraction methods. The highest accuracy rates are

also reported in recent papers and calculated the percentage of each model employed in these investigations.

Data Availability Statement: Upon reasonable request, the North Shore Health Institute in Chicago will provide the data supporting the study's conclusions.

Source of Funding: This study was not supported by any funding.

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethics Statement: The dataset for this investigation was obtained from the NorthShore Institute of Health System. The NorthShore University Health System Institutional Review Board examined and authorized the studies that used human subjects in the original collection of this dataset. This study did not require written informed consent from patients/participants for the use of this dataset, in compliance with national legislation and institutional standards.

REFERENCES

- [1] J. Parkinson, "An essay on the shaking palsy. 1817.," *J. Neuropsychiatry Clin. Neurosci.*, vol. 14, no. 2, Mar. 2002, doi: 10.1176/JNP.14.2.223.
- [2] A. Lees, "An essay on the shaking palsy," *Brain*, vol. 140, no. 3, pp. 843–848, 2017, doi: 10.1093/brain/awx035.
- [3] J. E. Ahlskog and M. D. Muentner, "Frequency of levodopa-related dyskinesias and motor fluctuations as estimated from the cumulative literature," *Mov. Disord.*, vol. 16, no. 3, pp. 448–458, May 2001, doi: 10.1002/MDS.1090.
- [4] Q. Zhang, G. M. Aldridge, N. S. Narayanan, S. W. Anderson, and E. Y. Uc, "Approach to Cognitive Impairment in Parkinson's Disease," *Neurotherapeutics*, vol. 17, no. 4, pp. 1495–1510, Oct. 2020, doi: 10.1007/S13311-020-00963-X.
- [5] Elliot S. Krames, P. Hunter Peckham, Ali R. Rezai, "Neuromodulation," Academic Press, 2009, Pages 1045-1063, ISBN 9780123742483. <https://doi.org/10.1016/B978-0-12-374248-3.00103-8>.
- [6] M. F. Contarino et al., "Cognitive outcome 5 years after bilateral chronic stimulation of subthalamic nucleus in patients with Parkinson's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 78, no. 3, pp. 248–252, Mar. 2007, doi: 10.1136/JNNP.2005.086660.
- [7] F. M. Weaver et al., "Bilateral deep brain stimulation vs best medical therapy for patients with advanced parkinson disease: A randomized controlled trial," *JAMA - J. Am. Med. Assoc.*, vol. 301, no. 1, pp. 63–73, Jan. 2009, doi: 10.1001/jama.2008.929.
- [8] E. Tripoliti et al., "Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease," *Neurology*, vol. 76, no. 1, pp. 80–86, Jan. 2011, doi: 10.1212/WNL.0B013E318203E7D0.
- [9] K. Witt et al., "Relation of lead trajectory and electrode position to neuropsychological outcomes of subthalamic neurostimulation in Parkinson's disease: results from a randomized trial," *Brain*, vol. 136, no. 7, pp. 2109–2119, Jul. 2013, doi: 10.1093/BRAIN/AWT151.
- [10] C. Oedekoven, L. Egeri, F. Jessen, M. Wagner, and R. Dodel, "Subjective cognitive decline in idiopathic Parkinson's disease: A systematic review," *Ageing Res. Rev.*, vol. 74, Feb. 2022, doi: 10.1016/J.ARR.2021.101508.
- [11] D. Aarsland et al., "Mild cognitive impairment in Parkinson disease: a multicenter pooled analysis," *Neurology*, vol. 75, no. 12, pp. 1062–1069, Sep. 2010, doi: 10.1212/WNL.0B013E3181F39D0E.
- [12] M. Emre et al., "Clinical diagnostic criteria for dementia associated with Parkinson's disease," *Mov. Disord.*, vol. 22, no. 12, pp. 1689–1707, Sep. 2007, doi: 10.1002/MDS.21507.
- [13] R. Monastero et al., "Mild cognitive impairment in Parkinson's disease: the Parkinson's disease cognitive study (PACOS)," *J. Neurol.*, vol. 265, no. 5, pp. 1050–1058, May 2018, doi: 10.1007/S00415-018-8800-4.
- [14] R. Morris et al., "Gait Rather Than Cognition Predicts Decline in Specific Cognitive Domains in Early Parkinson's Disease," *J. Gerontol. A. Biol. Sci. Med. Sci.*, vol. 72, no. 12, pp. 1656–1662, Nov. 2017, doi: 10.1093/GERONA/GLX071.
- [15] K. L. Wood et al., "Different PD-MCI criteria and risk of dementia in Parkinson's disease: 4-year longitudinal study," *npj Park. Dis.*, vol. 2, no. 1, pp. 1–8, Jan. 2016, doi: 10.1038/npjparkd.2015.27.
- [16] R. B. Postuma et al., "MDS clinical diagnostic criteria for Parkinson's disease," *undefined*, vol. 30, no. 12, pp. 1591–1601, Oct. 2015, doi: 10.1002/MDS.26424.
- [17] K. E. Stephan and K. J. Friston, "Functional Connectivity," *Encycl. Neurosci.*, pp. 391–397, 2009, doi: 10.1016/B978-008045046-9.00308-9.
- [18] M. Teplan, "Fundamentals Of EEG Measurement," *Meas. Sci. Rev.*, vol. 2, no. 2, pp. 1–11, 2003, [Online]. Available: papers3://publication/uuid/41BC5DF1-2B37-4955-BA95-42DF7DA40513
- [19] M. Akin, "Comparison of Wavelet Transform and FFT Methods in the Analysis of EEG Signals," *J. Med. Syst.* 2002 263, vol. 26, no. 3, pp. 241–247, Jun. 2002, doi: 10.1023/A:1015075101937.
- [20] C. Babiloni et al., "Abnormalities of resting-state functional cortical connectivity in patients with dementia due to Alzheimer's and Lewy body diseases: an EEG study," *Neurobiol. Aging*, vol. 65, pp. 18–40, May 2018, doi: 10.1016/j.neurobiolaging.2017.12.023.
- [21] C. Babiloni et al., "Abnormalities of Resting State Cortical EEG Rhythms in Subjects with Mild Cognitive

- Impairment Due to Alzheimer's and Lewy Body Diseases," *J. Alzheimers. Dis.*, vol. 62, no. 1, pp. 247–268, 2018, doi: 10.3233/JAD-170703.
- [22] K. T. E. Olde Dubbelink et al., "Resting-state functional connectivity as a marker of disease progression in Parkinson's disease: A longitudinal MEG study," *NeuroImage Clin.*, vol. 2, no. 1, pp. 612–619, Jan. 2013, doi: 10.1016/J.NICL.2013.04.003.
- [23] "Sci-Hub | Comparison of quantitative EEG between patients with Alzheimer's disease and those with Parkinson's disease dementia. *Clinical Neurophysiology*, 124(10), 1970–1974 | 10.1016/j.clinph.2013.05.001." <https://sci-hub.se/10.1016/j.clinph.2013.05.001> (accessed Nov. 06, 2022).
- [24] J. N. Caviness et al., "Both early and late cognitive dysfunction affects the electroencephalogram in Parkinson's disease," *Parkinsonism Relat. Disord.*, vol. 13, no. 6, pp. 348–354, Aug. 2007, doi: 10.1016/J.PARKRELDIS.2007.01.003.
- [25] V. J. Geraedts et al., "Machine learning for automated EEG-based biomarkers of cognitive impairment during Deep Brain Stimulation screening in patients with Parkinson's Disease," *Clin. Neurophysiol.*, vol. 132, no. 5, pp. 1041–1048, May 2021, doi: 10.1016/j.clinph.2021.01.021.
- [26] E. O. Brigham, "The Fast Fourier Transform and Its Applications," 1988, Accessed: Aug. 04, 2022. [Online]. Available: http://sar.kangwon.ac.kr/gisg/FFT_book.pdf
- [27] M. Akin, "Comparison of Wavelet Transform and FFT Methods in the Analysis of EEG Signals," *J. Med. Syst.* 2002 263, vol. 26, no. 3, pp. 241–247, Jun. 2002, doi: 10.1023/A:1015075101937.
- [28] M. Murugappan, S. Murugappan, B. Balaganapathy, and C. Gerard, "Wireless EEG signals based Neuromarketing system using Fast Fourier Transform (FFT)," *Proc. - 2014 IEEE 10th Int. Colloq. Signal Process. Its Appl. CSPA 2014*, pp. 25–30, 2014, doi: 10.1109/CSPA.2014.6805714.
- [29] L. M. Patnaik and O. K. Manyam, "Epileptic EEG detection using neural networks and post-classification," *Comput. Methods Programs Biomed.*, vol. 91, no. 2, pp. 100–109, Aug. 2008, doi: 10.1016/J.CMPB.2008.02.005.
- [30] U. Seljuq, F. Himayun, and H. Rasheed, "Selection of an optimal mother wavelet basis function for ECG signal denoising," *17th IEEE Int. Multi Top. Conf. Collab. Sustain. Dev. Technol. IEEE INMIC 2014 - Proc.*, pp. 26–30, 2014, doi: 10.1109/INMIC.2014.7096905.
- [31] B. Citation Venkata Phanikrishna, P. J. Pławiak, and A. Prakash, "A Brief Review on EEG Signal Pre-processing Techniques for Real-Time Brain-Computer Interface Applications," no. March 2022, 2021, doi: 10.36227/techrxiv.16691605.
- [32] M. Coles, & M. Rugg, "Event-related brain potentials", (1996). https://www.researchgate.net/publication/200003678_Event-related_brain_potentials.
- [33] J. Pardey, S. Roberts, and L. Tarassenko, "A review of parametric modelling techniques for EEG analysis," *Med. Eng. Phys.*, vol. 18, no. 1, pp. 2–11, 1996, doi: 10.1016/1350-4533(95)00024-0.
- [34] S. Sur and V. Sinha, "Event-related potential: An overview," *Ind. Psychiatry J.*, vol. 18, no. 1, p. 70, 2009, doi: 10.4103/0972-6748.57865.
- [35] D. E. J. Linden, "The p300: where in the brain is it produced and what does it tell us?," *Neuroscientist*, vol. 11, no. 6, pp. 563–576, Dec. 2005, doi: 10.1177/1073858405280524.
- [36] M. Chaturvedi et al., "Phase lag index and spectral power as QEEG features for identification of patients with mild cognitive impairment in Parkinson's disease," *Clin. Neurophysiol.*, vol. 130, no. 10, pp. 1937–1944, 2019, doi: 10.1016/j.clinph.2019.07.017.
- [37] V. J. Geraedts et al., "Clinical correlates of quantitative EEG in Parkinson disease: A systematic review," *Neurology*, vol. 91, no. 19, pp. 871–883, 2018, doi: 10.1212/WNL.0000000000006473.
- [38] H. M. M. Smeding, J. D. Speelman, H. M. Huizenga, P. R. Schuurman, and B. Schmand, "Predictors of cognitive and psychosocial outcome after STN DBS in Parkinson's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 82, no. 7, pp. 754–760, 2011, doi: 10.1136/jnnp.2007.140012.
- [39] N. Kumar, K. Khaund, and S. M. Hazarika, "Bispectral Analysis of EEG for Emotion Recognition," in *Procedia Computer Science*, 2016, vol. 84, pp. 31–35. doi: 10.1016/j.procs.2016.04.062.
- [40] V. J. Geraedts, M. L. Kuijf, J. J. van Hilten, J. Marinus, M. Oosterloo, and M. F. Contarino, "Selecting candidates for Deep Brain Stimulation in Parkinson's disease: the role of patients' expectations," *Park. Relat. Disord.*, vol. 66, no. December 2018, pp. 207–211, 2019, doi: 10.1016/j.parkreldis.2019.07.011.
- [41] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques," pp. 1–27, 2023.
- [42] T. Dabbabi, L. Bouaffif, and A. Cherif, "A Review of Non-Invasive Methods of Brain Activity Measurements via EEG Signals Analysis," *2023 IEEE Int. Conf. Adv. Syst. Emergent Technol.*, pp. 1–6, 2023.
- [43] J. Alexander, et al., "Resting functional connectivity and mild cognitive impairment in Parkinson's disease. An electroencephalogram study", *Future Neurol.*(2019) FNL18eISSN 1748-697110.2217/fnl-2018-0048C©2019 Jairo Alexander Carmona Arroyave, Jairo Alexander Carmona Arroyave, 2019.
- [44] Ponsen, M & Stam, C & Bosboom, J & Berendse, Henk & Hillebrand, Arjan, "A three dimensional anatomical view of oscillatory resting-state activity and functional connectivity in Parkinson's disease related dementia: An MEG study using atlas-based beamforming. *NeuroImage : clinical*", accessed Aug. 03, 2022. 2. 95-102. 10.1016/j.nicl.2012.11.007.
- [45] K. E. Stephan and K. J. Friston, "Functional connectivity," Stephan, K E; Friston, K J (2009). *Funct. Connect.*

- Squire, L R. *Encycl. Neurosci.* New York Elsevier, 391-397., pp. 391–397, 2009, doi: 10.1016/B978- 008045046-9.00308-9.
- [46] D. Stoffers, J. L. W. Bosboom, J. B. Deijen, E. C. Wolters, C. J. Stam, and H. W. Berendse, “Increased cortico-cortical functional connectivity in early-stage Parkinson’s disease: a MEG study,” *Neuroimage*, vol. 41, no. 2, pp. 212–222, Jun. 2008, doi: 10.1016/J.NEUROIMAGE.2008.02.027.
- [47] L. I. Boon, A. Hillebrand, K. T. E. Olde Dubbelink, C. J. Stam, and H. W. Berendse, “Changes in resting-state directed connectivity in cortico-subcortical networks correlate with cognitive function in Parkinson’s disease,” *Clin. Neurophysiol.*, vol. 128, no. 7, pp. 1319–1326, Jul. 2017, doi: 10.1016/J.CLINPH.2017.04.024.
- [48] V. J. Geraedts et al., “Quantitative EEG reflects non-dopaminergic disease severity in Parkinson’s disease,” *Clin. Neurophysiol.*, vol. 129, no. 8, pp. 1748–1755, 2018, doi: 10.1016/j.clinph.2018.04.752.
- [49] R. L. Utianski et al., “Graph theory network function in parkinson’s disease assessed with electroencephalography,” *Clin. Neurophysiol.*, vol. 127, no. 5, pp. 2228–2236, 2016, doi: 10.1016/j.clinph.2016.02.017.
- [50] L. Bonanni, “The democratic aspect of machine learning: limitations and opportunities for Parkinson’s disease”, *Mov Disord* 2019; 34(2):164–6.
- [51] M. B. Kursu and W. R. Rudnicki, “Feature selection with the boruta package,” *J. Stat. Softw.*, vol. 36, no. 11, pp. 1–13, 2010, doi: 10.18637/jss.v036.i11.
- [52] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, “Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package),” *Neurocomputing*, vol. 307, pp. 72–77, 2018, doi: 10.1016/j.neucom.2018.03.067.
- [53] T. H. Grandy, M. Werkle-Bergner, C. Chicherio, F. Schmiedek, M. Lövdén, and U. Lindenberger, “Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults,” *Psychophysiology*, vol. 50, no. 6, pp. 570–582, Jun. 2013, doi: 10.1111/PSYP.12043.
- [54] A. Goljahani, C. D’Avanzo, S. Schiff, P. Amodio, P. Bisiacchi, and G. Sparacino, “A novel method for the determination of the EEG individual alpha frequency,” *Neuroimage*, vol. 60, no. 1, pp. 774–786, Mar. 2012, doi: 10.1016/J.NEUROIMAGE.2011.12.001
- [55] A. M. Maitín, A. J. García-Tejedor, and J. P. R. Muñoz, “Machine learning approaches for detecting parkinson’s disease from eeg analysis: A systematic review,” *Appl. Sci.*, vol. 10, no. 23, pp. 1–21, 2020, doi: 10.3390/app10238662.
- [56] Y. Si, “Machine learning applications for electroencephalograph signals in epilepsy: A quick review,” *Acta Epileptol.*, vol. 2, no. 1, pp. 5–11, 2020, doi: 10.1186/s42494-020-00014-0.
- [57] V. Doma and M. Pirouz, “A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals,” *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00289- 7.
- [58] A. Pal, M. Behari, V. Goyal, and R. Sharma, “Study of EEG microstates in Parkinson’s disease: a potential biomarker?” *Cogn. Neurodyn.*, vol. 15, no. 3, pp. 463–471, 2021, doi: 10.1007/s11571-020-09643-0.
- [59] A. Joshi, A. Kamble, A. Parate, S. Parkar, D. Puri, and C. Gaikwad, “Drowsiness Detection using EEG signals and Machine Learning Algorithms,” *ITM Web Conf.*, vol. 44, p. 03030, 2022, doi: 10.1051/itmconf/20224403030.
- [60] Hospital Health System in the Chicago Area | NorthShore [Internet]. [cited 2023 Jun 29]. Available from: <https://www.northshore.org/>
- [61] G. Dautzenberg, J. Lijmer, and A. Beekman, “Clinical value of the Montreal Cognitive Assessment (MoCA) in patients suspected of cognitive impairment in old age psychiatry. Using the MoCA for triaging to a memory clinic,” *Cogn. Neuropsychiatry*, vol. 26, no. 1, pp. 1–17, 2021, doi: 10.1080/13546805.2020.1850434.
- [62] “MoCA test for dementia: Results, diagnosis, and more.” <https://www.medicalnewstoday.com/articles/moca- test-for-dementia#uses> (accessed Mar. 24, 2023).
- [63] V. J. Geraedts et al., “Preoperative Electroencephalography-Based Machine Learning Predicts Cognitive Deterioration After Subthalamic Deep Brain Stimulation,” *Mov. Disord.*, vol. 36, no. 10, pp. 2324–2334, 2021, doi: 10.1002/mds.28661.
- [64] K. Novak, B. A. Chase, J. Narayanan, P. Indic, and K. Markopoulou, “Quantitative Electroencephalography as a Biomarker for Cognitive Dysfunction in Parkinson’s Disease,” *Front. Aging Neurosci.*, vol. 13, no. January, 2022, doi: 10.3389/fnagi.2021.804991.
- [65] V. V. Cozac, U. Gschwandtner, F. Hatz, M. Hardmeier, S. Rüegg, and P. Fuhr, “Quantitative EEG and cognitive decline in Parkinson’s disease,” *Parkinsons. Dis.*, vol. 2016, 2016, doi: 10.1155/2016/9060649.
- [66] I. Litvan et al., “MDS Task Force on mild cognitive impairment in Parkinson’s disease: critical review of PD-MCI,” *Mov. Disord.*, vol. 26, no. 10, pp. 1814–1824, Aug. 2011, doi: 10.1002/MDS.23823.
- [67] E. Arasteh, A. Mahdizadeh, M. S. Mirian, S. Lee, and M. J. McKeown, “Deep Transfer Learning for Parkinson’s Disease Monitoring by Image-Based Representation of Resting-State EEG Using Directional Connectivity,” 2021, doi: 10.3390/a15010005.
- [68] L. Molcho, N. B. Maimon, N. Regev-Plotnik, S. Rabinowicz, N. Intrator, and A. Sasson, “Single-Channel EEG Features Reveal an Association with Cognitive Decline in Seniors Performing Auditory Cognitive Assessment,” *Front. Aging Neurosci.*, vol. 14, no. May, pp. 1–19, 2022, doi: 10.3389/fnagi.2022.773692.
- [69] M. Kopańska, D. Ochojska, A. Dejniewicz-Velitchkov, and A. Banaś-Ząbczyk, “Quantitative Electroencephalography (QEEG) as an Innovative Diagnostic Tool in Mental Disorders,” *Int. J. Environ. Res.*

Public Health, vol. 19, no. 4, 2022, doi: 10.3390/ijerph19042465.

[70] S. Chatzikonstantinou, J. McKenna, E. Karantali, F. Petridis, D. Kazis, and I. Mavroudis, "Electroencephalogram in dementia with Lewy bodies: a systematic review," *Aging Clin. Exp. Res.*, vol. 33, no. 5, pp. 1197–1208, May 2021, doi: 10.1007/S40520-020-01576-2.

[71] M. Cai et al., "Identifying Mild Cognitive Impairment in Parkinson's Disease with Electroencephalogram Functional Connectivity," *Front. Aging Neurosci.*, vol. 13, Jul. 2021, doi: 10.3389/fnagi.2021.701499.

[72] U. Gschwandtner, G. Bogaarts, V. Roth, and P. Fuhr, "Prediction of cognitive decline in Parkinson's disease (PD) patients with electroencephalography (EEG) connectivity characterized by time-between-phase-crossing (TBPC)," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-32345-6.

[73] Z. Li, K. Hwang, K. Li, J. Wu, and T. Ji, "Graph - generative neural network for EEG - based epileptic seizure detection via discovery of dynamic brain functional connectivity," *Sci. Rep.*, pp. 1–16, 2022, doi: 10.1038/s41598-022-23656-1.

[74] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, "EEG-GNN: Graph Neural Networks for Classification of Electroencephalogram (EEG) Signals," Jun. 2021, [Online]. Available: <http://arxiv.org/abs/2106.09135>

[75] P. Zhong, D. Wang, and C. Miao, "EEG-Based Emotion Recognition Using Regularized Graph Neural Networks," Jul. 2019, [Online]. Available: <http://arxiv.org/abs/1907.07835>

[76] A. E. Lang et al., "Deep brain stimulation: Preoperative issues," *Mov. Disord.*, vol. 21, no. SUPPL. 14, 2006, doi: 10.1002/mds.20955.

[77] V. J. Geraedts et al., "Clinical correlates of quantitative EEG in Parkinson disease," pp. 1–14, 2018, doi: 10.1212/WNL.0000000000006473.

[78] J. Harvey et al., "Machine learning-based prediction of cognitive outcomes in de novo Parkinson's disease," *npj Park. Dis.*, vol. 8, no. 1, 2022, doi: 10.1038/s41531-022-00409-5.

[79] E. Uc, F. Anjum, S. Dasgupta, and N. Narayanan, "Resting-state EEG Predicts Cognitive Impairment in Parkinson's Disease (P6-11.015)," *Neurology*, vol. 100, no. 17_supplement_2, 2023, doi: 10.1212/wnl.0000000000203685.

[80] F. Anjum et al., "Resting-state EEG measures cognitive impairment in Parkinson's disease," Preprint, pp. 1–21, 2023.

[81] M. Parajuli, A. W. Amara, and M. Shaban, "Screening of Mild Cognitive Impairment in Patients with Parkinson's Disease Using a Variational Mode Decomposition Based Deep-Learning," in *International IEEE/EMBS Conference on Neural Engineering, NER*, 2023. doi: 10.1109/NER52421.2023.10123759.

[82] M. Parajuli, A. W. Amara, and M. Shaban, "A Novel Deep-Learning Based Approach for Mild Cognitive Impairment Screening in Patients with Parkinson's Disease," in *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things, AIBThings 2023 - Proceedings*, 2023. doi: 10.1109/AIBThings58340.2023.10292456.

[83] M. Parajuli, A. W. Amara, and M. Shaban, "Deep-learning detection of mild cognitive impairment from sleep electroencephalography for patients with Parkinson's disease," *PLoS One*, vol. 18, no. 8 August, 2023, doi: 10.1371/journal.pone.0286506.

[84] M. B. Segura, J. Bray, J. Gagnon, and B. Platt, "Novel EEG tools to predict early cognitive impairments in Parkinson's disease," *Alzheimer's Dement.*, vol. 19, no. S15, 2023, doi: 10.1002/alz.076164.

[85] H. Chang, B. Liu, Y. Zong, C. Lu, and X. Wang, "EEG-Based Parkinson's Disease Recognition via Attention-Based Sparse Graph Convolutional Neural Network," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 11, 2023, doi: 10.1109/JBHI.2023.3292452.

[86] M. Amboni et al., "Resting-state functional connectivity associated with mild cognitive impairment in Parkinson's disease," *J. Neurol.*, vol. 262, no. 2, 2015, doi: 10.1007/s00415-014-7591-5.

[87] R. Monastero et al., "Mild cognitive impairment in Parkinson's disease: the Parkinson's disease cognitive study (PACOS)," *J. Neurol.*, vol. 265, no. 5, 2018, doi: 10.1007/s00415-018-8800-4.