|| Jai Sri Gurudev ||

Sri Adichunchanagiri Shikshana Trust®

ADICHUNCHANAGIRI INSTITUTE OF TECHNOLOGY

Jyothi Nagara, Chikkamagaluru-577102, Karnataka, India.

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Second IEEE International Conference on Advances in Information Technology (ICAIT-24)

IEEE

ICAIT

Technical Co-Sponsorer IEEE BANGALORE SECTION

Certificate of Appreciation

This is to certify that Dr./Prof./Mr./Mrs./Ms. **Malkeet Singh** from **Graphic Era Hill University, Dehradun** has participated/presented a paper titled **Parkinson's Disease Detection and Analysis Using Machine Learning** in the **Second IEEE International Conference on Advances in Information Technology (ICAIT-24)** Organized by the Department of Computer Science & Engineering, AIT, Chikkamagaluru, Karnataka, India during **24th-27th July 2024**.

Dr. Puspha Ravikumar
Prof. & Head, Dept. of CS&E
Organizing Chair, ICAIT-24

Dr. C.T Jayadeva
Principal, AIT, Chikkamagaluru
General Chair, ICAIT-24

Dr. C.K Subbaraya
Director, AIT & Registrar, ACU
Patron, ICAIT-24

Parkinson's Disease Detection and Analysis Using Machine Learning

Aniruddha Prabhu B P
Department of Computer Science and Engineering
Alliance University
Bangalore, India
aniprabhubp@gmail.com

Chandradeep Bhatt
Department of Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
bhattchandradeep@gmail.com

Noor Mohd
Department of Computer Science and Engineering
Graphic Era (Deemed to be) University
Dehradun, India
noormohd@geu.ac.in

Tushar Sharma
Department of Computer Science and Engineering
Graphic Era (Deemed to be) University
Dehradun, India
tusharsharma0804@gmail.com

Malkeet Singh
Department of Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
malkeetchahal34@gmail.com

***Abstract*— Neurodegenerative disease Parkinson's, which have both motor and non-motor type of symptoms, is one of the biggest problems, which gets consideration in the public health department. To assurance of patients and beneficial therapeutics process diagnosis should occur in the critical stages and should also be precise. The aim of this study is developing of the machine learning methods for the diagnosis and appraisal of Parkinson's disease. Dataset which contains patient data is used where each record has important attributes such as age, height, weight, gender, and so forth is obtained from the University of California Irvine (UCI) Machine Learning Repository. The research process commences with a detailed data-gathering procedure which is followed by an investigative exploratory data analysis to discover an inherent pattern shared by all Parkinsonian traits. The data visualization tools can help us more deep information about Parkinson's disease symptoms. The task of making sure that the Minority Random Oversampling method is properly implemented in data preparation is quite significant for us as we work on the project because it is one of the important aspects. On the other hand, the flare Light Gradient Boosted Machine (LightGBM) model is the one we use for this particular purpose. When run on the test dataset, this state of the art algorithm hits the almost perfect result of 99.24% accuracy. As well as adding up new facts to the science of the Parkinson's disease research, this work shows how powerful computer algorithms and perfect datasets are in the spreading accurate diagnoses of Parkinson's disease.**

Keywords*— parkinson, neurology, EEG, LightGBM, Health, Data Analysis.

I. Introduction

Parkinson's Disease (PD) is an advancing degeneration of the nervous system that the outcomes of which are considered very distressing to many people and their families across the globe. The way that clinical features of Parkinson's Disease change is quite unpredictable for doctors, consisting of different categories of motor and non-motor complications. Ranging from light shaking to inability to move normally, all kinds of them take the quality life down to the bottom and creativity of physicians up from the bottom. With the estimation of about 10 million people on the global scoped population (World Health Organization, 2019), there is a great need for a better diagnosis method for diagnosis to be more accurate and also time saving.

Traditional diagnostic techniques of Parkinson's disease employ clinical examination and subjective criteria for assessing symptoms, a process likely to be limited in terms of precision, speed, and the role of the human factor as an obstacle. After meeting these challenges, the combination of the modern technologies came as a prospective solution. Thus, there is a growing reliance on special algorithms, mainly machine learning, both for the enhancing of diagnostic accuracy and accelerating personalized therapy.

The union of Machine Learning and healthcare makes a significant breakthrough in finding solutions for the tough problems of Parkinson's disease that originate from Neurodegeneration. Machine learning algorithms which are known as being proficient in finding complicated patterns within quite large data sets are strong tools that can be used for diagnosis, prognosis, and planning of the treatment. The use of machine learning algorithms can be considered as an alternative form of detection in Parkinson's disease screening which is more efficient than the traditional methodologies. Employing the tangible data obtained from brain scan machine learning model can be trained to recognize brain patterns which are unique to Parkinson's Disease and be able to use the distinctive characteristics to identify Parkinson's Disease successfully and timely.

This studies states on how machine learning tools are used to diagnose and assess Parkinson's Disease using a dataset provided from the University of California Irvine (UCI) Machine Learning repository. The collection portrays the complexity of the disease - each file comprised of a 23.6 second-long recording of brain activity. The problem of unbalanced data is realized through the Minority Random Oversampling approach taken during preprocessing which will even the representation of different brain functions. EDA (exploratory data analysis) shows hidden patterns of the dataset and gives us some temporal facts which happen during Parkinsonian episodes. We construct a data visualization model using the properties of Parkinsonian brain activity which laid the basis for the following feature selection and model development.

Several models like decision trees, support vector machines, and gradient boosting are applied for the symptoms of the disease forecasting and analyzing of Parkinson's. Among these systems, the Light Gradient Boosted Machine (LightGBM) classifier algorithm constantly achieves outstanding performance, which is characterized by quick forecasts and useful patterns. Providing an excellent performance, it is based on the machine learning algorithm, which can cope with the data of any scale and proves to be effective.

The results of our research propose a fully personalized approach to the Parkinson disease detection addressing the dataset problems and incorporating temporal information found in neural signals. Through the comparison of machine learning capabilities, especially the LightGBM model, the study sets an example not only for Parkinson's Disease research but also highlights the impact of balanced sets and advanced models in improving machine learning accuracy and efficacy in disease detection. In short, our findings add to the knowledge of PD and indicate the central role of complex machine learning algorithms which can boost the diagnostic skills.

II. Literature Survey

Parkinson's disease is a common neurological condition which affects thousands upon thousands around the world and this is characterized as a result of its unpredictable behavior. There exists the ability for early identification of Parkinson's symptoms and superb prognosis which can transform the treatment of the disease. Over the last several years, machine learning that can be connected to diverse data sources has been a potential solution towards these pursuits. This analysis of literature has a clear objective to evaluate the fast-growing group of studies that use machine learning models for diagnoses and prediction of Parkinson's disease. Through reviewing various studies, methodologies, datasets, and consequences, we try to present an in-depth overall picture of the here and now in the terms of the neurology and artificial intelligence connection.

Bravo-Vázquez and colleagues (2023) have provided a review of the FDA-approved therapies and the clinical trials' trends for the Alzheimer's and Parkinson's diseases. Instead of being given, it underlines the concept of counter-accumulation, anti-inflammation, and regeneration revealing into the transforming medicine arena. The study highlights AD and PD therapeutic strategy to divulge in future and underutilization of delivery vehicles indicating how the landscape of biologic therapies has been changing [1] Alnaaim et al., (2023) Alluded to the role of the Liver X Receptor (LXR) in Neuroprotective Parkinson's disease (PD) against the neuroinflammation, oxidative stress, and mitochondrial dysfunction. Changing LXR seems to be a promising way that may limit, or possibly even prevent the neurodegeneration in PD [2]. Not only did Ayaz et al. (2023) cleave but also cut artificial intelligence applications for Parkinson's disease diagnosis across a heterogeneous dataset with versatile machine learning techniques. The review highlights how the tools have the potential to bring accuracy and efficiency in the area of disease diagnostics and this addresses one of the major challenges in the field. Haider et al. (2023) demonstrate the breakthroughs in molecular imaging, precisely, these probes includes α-synuclein, mitochodrial dysfunction, and neuroinflammation for Parkinson's disease. This review offers information on imaging modalities and other peripheral markers, which can be used to seek Researches about diagnostic tracers and therapeutic interventions [4].

Batiha et al. (2023) delve into Silent Information Regulator 1's (SIRT1) neuroprotective effects in Parkinson's disease, highlighting its role in mitigating oxidative stress, inflammation, and neuronal cell death. The review positions SIRT1 as a crucial player in addressing PD neuropathology [5]. Ali and Chakraborty propose a novel ensemble approach named EOFSC for Parkinson's disease detection. Integrates feature selection with a deep neural network, outperforming conventional models and enhancing accuracy by 6.5% [6]. Xu et al. analyzes bidirectional interactions among microglia and T lymphocytes in Parkinson's disease, offering information on their impact on neuroinflammation. Highlights potential targets and models for future research directions [7]. Bravo-Vázquez et al. reviews current advances in plant-based vaccines for neurodegenerative diseases, emphasizing the evolving therapeutic strategies for Alzheimer's and Parkinson's diseases [8]. Zhu et al. gives a detailed analysis of the immunobiology of Parkinson's illness, concentrating on α-synuclein's position in the gut-brain connection theory and the participation of both adaptive and innate immune responses [9]. Spinelli and Haigis analyze the varied contributions made by mitochondria to cellular metabolism, stressing the function of mitochondrial quality-assurance mechanisms and mitophagy as to preserve cellular homeostasis [10].

Alarcón et al. studies the brain mechanisms behind the neuroprotection effects of environmental enrichment in Parkinson's disease. The study, part of the Park-in-Shape experiment, indicates that aerobic exercise maintains disease development in the corticostriatal sensory network and boosts cognitive function [11, 12]. Willis et al. evaluates the prevalence of Parkinson's disease in North America using data from five epidemiologic cohorts. Estimations of age-sex-adjusted PD prevalence reveal heterogeneity and emphasize the necessity of more accurate predictions for care planning and research [13, 14]. Heidari et al. examines the involvement of Toll-like receptors and neuroinflammation in Parkinson's disease. Highlights evidence of neuroinflammation, elevated TLR expression, and their potential role to dopaminergic neuronal degeneration in PD patients [15]. Espay and Lang explores the shift from a clinico-pathologic convergence model of Parkinson's disease to a systems biology divergence model. Emphasizes the necessity for subtype-specific therapy methods based on divergent, systems-biology models [16]. McFarthing provides an overview of pharmacological therapies in clinical trials for Parkinson's disease in 2021-2022. Categorizes therapies into symptomatic treatments (ST) and disease-modifying treatments (DMT) and discusses the progress made despite the ongoing COVID-19 pandemic [17].

Mestre et al. critically evaluates Parkinson's disease (PD) subtyping systems through a systematic review of 38 studies. The review identifies methodologic shortcomings, questionable clinical applicability, and unknown biological relevance, suggesting the need for new approaches acknowledging individual-level heterogeneity [18]. Fanning et al. explores the role of lipids in Parkinson's disease, proposing a "lipid cascade." Discusses the association of the protein α-synuclein with PD and other synucleinopathies and emphasizes the significance of deciphering the biochemistry of α-synuclein in native systems for treatment development [19]. Zhao et al analyzes the mutational spectrum of familial Parkinson's disease and sporadic early-onset Parkinson's disease in a mainland Chinese population. Identifies harmful mutations in known Parkinson's disease-associated genes, underscoring the necessity of genetic testing, especially in early-onset patients [20]. Jun and Kim explain the MDS Research Criteria for Prodromal Parkinson's Disease, combining new data and indicators. Utilizes a Bayesian classifier technique, accounting for age and predictive information from risk and prodromal indicators to assess the likelihood of prodromal PD [21]. Ticinesi et al. discusses the

potential role of gut microbiota as a mediator of the beneficial effects of dietary (poly) phenols on skeletal muscle in aging. While the abstract provided does not directly discuss Parkinson's disease, it points to the growing recognition of the gut-brain axis in various neurological disorders, including PD [22].

III. Proposed Methodology

The following methodology is a holistic approach employing the latest machine learning technology in the effort to detect the disease Parkinson's for both speediness, dependability and accuracy. The systematic procedure encompasses the crucial phases such as selecting data up to the creation and deployment of machine learning models.

A. Data Collection & Preprocessing

The first phase is about the data acquisition and their further processing in order to build the system of illness recognition on the solid base. The dataset has been taken from University of California, Irvine ML Repository which consists of around 100 files arranged in 5 folders. Every person's file is uniquely different and contains a 23.6-second recording of each person's brain activity. Sampling is done on the dataset with a data point of 4097, this capture EEG records at various moments. Consequently, the data points are grouped and partitioned into 23 pieces among which the second dataset will be formed, which will be a new data set of 11500 items, each including 178 data points for 1 second [Fig. 1]. The dataset has several activities which are portrayed by y that includes the various symptoms, the locations where the disease is in the spread, the healthy brain parts, and the different states such as resting and moving. This binary estimation is performed by means of y being transformed binary, where 1 means Parkinson and 0 means non-Parkinson [Fig. 2].

	Unnamed	X1	X2	X3	X4	X5	X6	X7	X8	X9	...	X170	X171	X172	X173	X174	X175	X176	X177	X178	y
0	X21.V1.791	135	190	229	223	192	125	55	-9	-33	...	-17	-15	-31	-77	-103	-127	-116	-83	-51	4
1	X15.V1.924	386	382	356	331	320	315	307	272	244	...	164	150	146	152	157	156	154	143	129	1
2	X8.V1.1	-32	-39	-47	-37	-32	-36	-57	-73	-85	...	57	64	48	19	-12	-30	-35	-35	-36	5
3	X16.V1.60	-105	-101	-96	-92	-89	-95	-102	-100	-87	...	-82	-81	-80	-77	-85	-77	-72	-69	-65	5
4	X20.V1.54	-9	-65	-98	-102	-78	-48	-16	0	-21	...	4	2	-12	-32	-41	-65	-83	-89	-73	5
5	X14.V1.56	55	28	18	16	16	19	25	40	52	...	-12	-31	-42	-54	-60	-64	-60	-56	-55	5
6	X3.V1.191	-55	-9	52	111	135	129	103	72	37	...	-125	-99	-79	-62	-41	-26	11	67	128	4
7	X11.V1.273	1	-2	-8	-11	-12	-17	-15	-16	-18	...	-79	-91	-97	-88	-76	-72	-66	-57	-39	2
8	X19.V1.874	-278	-246	-215	-191	-177	-167	-157	-139	-118	...	-400	-379	-336	-281	-226	-174	-125	-79	-40	1
9	X3.V1.491	8	15	13	3	-6	-8	-5	4	25	...	49	31	11	-5	-17	-19	-15	-15	-11	4
10	X3.V1.6	-5	15	28	28	9	-29	-41	-19	14	...	-38	-4	25	16	-16	-74	-101	-89	-49	5
11	X21.V1.724	-167	-230	-280	-315	-338	-369	-405	-392	-298	...	423	434	416	374	319	268	215	165	103	1
12	X7.V1.162	92	49	0	-32	-51	-65	-37	-19	-25	...	-56	-41	-40	-43	-32	-13	-1	-7	-44	3
13	X1.V1.211	15	12	0	-17	-28	-31	-39	-51	-44	...	-88	-102	-97	-77	-45	-19	13	44	68	4
14	X1.V1.615	-24	-15	-5	-1	4	3	6	10	11	...	32	35	36	34	32	26	23	18	20	2
15	X22.V1.242	-135	-133	-125	-118	-111	-105	-102	-93	-94	...	-49	-39	-35	-29	-10	4	21	31	37	3
16	X1.V1.863	39	41	41	42	43	43	46	47	49	...	43	41	41	43	43	40	41	41	49	2
17	X9.V1.302	9	4	-5	-10	-22	-30	-33	-43	-41	...	34	27	22	18	15	13	9	9	3	3
18	X7.V1.541	-21	-5	1	7	19	20	13	2	-1	...	43	28	25	19	30	35	26	5	-13	4
19	X9.V1.915	4	24	51	76	92	102	104	101	90	...	3	5	10	19	31	36	40	43	36	2

Fig. 1. Original EEG Dataset Showing Multi Classes for Parkinson

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	...	X170	X171	X172	X173	X174	X175	X176	X177	X178	y
0	135	190	229	223	192	125	55	-9	-33	-38	...	-17	-15	-31	-77	-103	-127	-116	-83	-51	0
1	386	382	356	331	320	315	307	272	244	232	...	164	150	146	152	157	156	154	143	129	1
2	-32	-39	-47	-37	-32	-36	-57	-73	-85	-94	...	57	64	48	19	-12	-30	-35	-35	-36	0
3	-105	-101	-96	-92	-89	-95	-102	-100	-87	-79	...	-82	-81	-80	-77	-85	-77	-72	-69	-65	0
4	-9	-65	-98	-102	-78	-48	-16	0	-21	-59	...	4	2	-12	-32	-41	-65	-83	-89	-73	0
5	55	28	18	16	16	19	25	40	52	66	...	-12	-31	-42	-54	-60	-64	-60	-56	-55	0
6	-55	-9	52	111	135	129	103	72	37	0	...	-125	-99	-79	-62	-41	-26	11	67	128	0
7	1	-2	-8	-11	-12	-17	-15	-16	-18	-17	...	-79	-91	-97	-88	-76	-72	-66	-57	-39	0
8	-278	-246	-215	-191	-177	-167	-157	-139	-118	-92	...	-400	-379	-336	-281	-226	-174	-125	-79	-40	1
9	8	15	13	3	-6	-8	-5	4	25	41	...	49	31	11	-5	-17	-19	-15	-15	-11	0
10	-5	15	28	28	9	-29	-41	-19	14	30	...	-38	-4	25	16	-16	-74	-101	-89	-49	0
11	-167	-230	-280	-315	-338	-369	-405	-392	-298	-140	...	423	434	416	374	319	268	215	165	103	1
12	92	49	0	-32	-51	-65	-37	-19	-25	-29	...	-56	-41	-40	-43	-32	-13	-1	-7	-44	0
13	15	12	0	-17	-28	-31	-39	-51	-44	-35	...	-88	-102	-97	-77	-45	-19	13	44	68	0
14	-24	-15	-5	-1	4	3	6	10	11	7	...	32	35	36	34	32	26	23	18	20	0
15	-135	-133	-125	-118	-111	-105	-102	-93	-94	-90	...	-49	-39	-35	-29	-10	4	21	31	37	0
16	39	41	41	42	43	43	46	47	49	50	...	43	41	41	43	43	40	41	41	49	0
17	9	4	-5	-10	-22	-30	-33	-43	-41	-40	...	34	27	22	18	15	13	9	9	3	0
18	-21	-5	1	7	19	20	13	2	-1	-3	...	43	28	25	19	30	35	26	5	-13	0

Fig. 2. EEG Dataset Showing Binary Classes for Parkinson

Addressing the imbalanced nature of the target variable y, the Minority Random Oversampling technique is applied during preprocessing. This intentional augmentation of cases from the minority class assures a balanced representation, boosting the model's capacity to recognize patterns associated with Parkinson's `Disease. This rigorous preprocessing technique examines temporal dynamics and imbalances, setting the basis for later exploratory data analysis and machine learning model development.

B. Exploratory Data Analysis and Visualization

Exploratory Data Analysis (EDA) are a very important part of the data science court as they enable discovering interesting patterns within the dataset. With a mind for the temporality of the disease, this type of plots has the ability to resolve details of brain activity fluctuations in the impulse recording period. With this high-resolution, intricate pattern, and markers can be spotted in the temporal EEG data, as potential sources of signals. A distinct advantage of the EDA produced by us is the generation of charts manifesting EEG patterns for Parkinsonian [Fig 3] and non-Parkinsonian [Fig 4] patients, giving a comprehensive visual picture for the purpose of identifying worn-out signs and new patterns necessary for the accurate determination of various brain states. Alongside the timeline sketches, frequency-based analysis is provided through spectrograms to complete the analysis. Spectrograms reveal interactions on spectra within EEG recordings, which assists to determine frequency patterns concerning with specific episodes of parkinsonian. This hybrid feature engineering tool takes the lead the job for other feature engineering efforts.

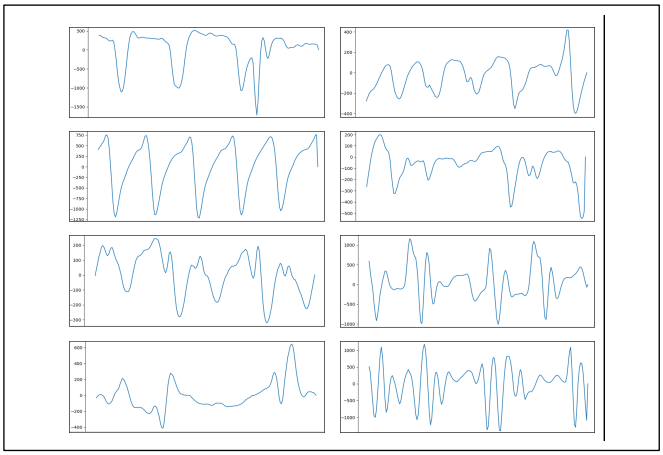

Fig. 3. EEG Plots for Parkinson Patients

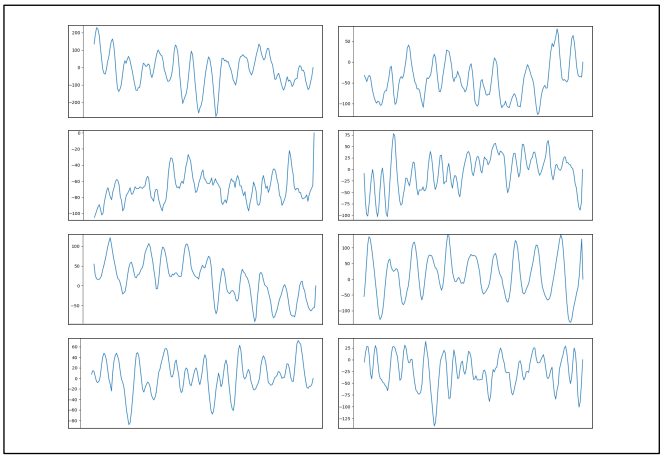

Fig. 4. EEG Plots for Non Parkinson Patients

C. *Splitting The Dataset For Training and Testing*

Data splitting into the training and testing parts is a fundamental step of the model development and verification process. The dataset is split systematically into the training part which consists 70% of the dataset and test part which contains the remaining 30% using train_test_split function. This Random Seed split runs consistently in the next analyses allowing for high reliability. Training set is the foundation of the model's learning environment providing the ability to identify immanent connections, influence regularities as well as the features. The set of test is the one that performs the function of the independent control, allowing to assess the predictive ability and generalizability to real-life situation.

D. *Model Selection*

The cruicial aspect of model selection is about choosing a machine learning model that keeps the changes in the EEG dataset and the subtleness of the Parkinson's Disease identification. Several models having statistical analysis of time series, ensembles of methods, and class average model are evaluated. By looking at the results of detailed testing, the ideal model of the Light Gradient Boosted Machine (LightGBM) becomes clear.Configured for binary classification, LightGBM leverages a gradient boosting framework, offering advantages such as parallel training and leaf-wise tree growth. Meticulous tuning of hyperparameters optimizes its performance, making it adept at handling imbalanced datasets, navigating high-dimensional feature spaces, and providing accurate predictions.

E. *Predictive Analysis*

The Predictive analysis phase marks the final stage, which is the training of the chosen LightGBM algorithm with a particular dataset subset and the thorough assessment of the performance of the system in Parkinson's disease detection. During this process, the network gets acquainted with the complexities of EEG records and can lean on the patterns and the relationship there. Tuned parameters control the process leading to the optimization, thus the model captures those features that characterize the presence and absence of the disease. This step involves applying the model to the reserved testing subset using metrics such as precision, specificity, sensitivity, and area under the Receiver Operating Characteristic curve (AUC-ROC) for a comprehensive performance check. Our holistic methodology is capable not only to use the most accurate way of detecting Parkinson's Disease but also to provide detailed information about the evolving state of the brain during the occurrence of such events.

IV. Results

The model built with the Light Gradient Boosting Machine (LightGBM) technology shows outstanding results in the framework of Parkinson's Disease detection, having a great accuracy which amounts to 99.24% on the test set. The confusion matrix [Fig 6] provides a detailed overview of the model's precision, accurately identifying Parkinsonian (TP: We explored working on sports related (TN: 2749) activities and non-Parkinson related (TN: 2754) ones while the reducing diagnosis (CN: 7) errors. Thus, the precise perpendicularity of these processes is a crucial point for the reliability of the condition evaluation of Parkinson's disease above all.

On the other hand, the analysis about the feature importance also reveals the significance of specific EEG signals, which are indispensable to make an accurate prediction, and thus it increased the interpretability of LightGBM model [Fig 7]. The LightGBM model's excellent accuracy, a correctly balanced confusion matrix, and revealing feature importance show the effectiveness of this model in discovering Parkinson's disease. The AUC-metric returns the maximum result of 0.91 for the model that is capable of classifying Parkinsonian tasks with high accuracy between Parkinsonian and non-Parkinsonian activities. In real word, these results indicate that the model shows the possibility of use as a trustworthy and understandable instrument for the application in the diagnosis or observation of the Parkinson's Disease. The high accuracy, low false positive and false negative rates, and valuable insights into the underlying dynamics of EEG recordings highlight the model's effectiveness in enhancing diagnostic precision for Parkinson's Disease.

V. Discussion & Future Works

The talk about, on the other hand, includes the elucidation of the implications of our results in the general level of PD diagnosis and the suggestion of what research we could do in the future to learn more about the disease. The excellent result of our Light Gradient Boosting Machine (LightGBM) model confirms it as one of the possible useful tools for practical applications in Parkinson's disease diagnosis, which shows its accuracy in the test with 99.24 percent and its curve of ROC robust with an AUC of 0.99. The fact that the model's fidelity in reducing false diagnosis of Parkinsonian's, as depicted in the confusion matrix, is really important in medical settings is the most vital part of the whole process. The high interpretability feature of our model via the feature importance analysis allows clinicians and researchers to isolate the EEG signals influencing predictions. This is yielded by the clear view offered regarding the decision-making process. Having such human readable nature is the reason for implementing machine learning models in medical applications.

Future research directions could be based on the extension of the dataset such that it mainly deals with people who are different from the dataset thus improving the models. On the other hand, looking into true implementation and ongoing surveillance may take our model to the next level of clinical implementation. Connecting the model with experts on

relevant fields and neurologists becomes crucial for improving its degree of interpretability and its ability to provide the results aligned with the clinical understanding. Our study is not only a robust model for Parkinson's Disease detection but also it is a gateway for researcher, and this created a concrete platform for collaboration in healthcare settings. Currently, the application of cutting-edge machine learning into expert frameworks is the field offering the most notable benefits to realize a revolution in the diagnosis of as well as care of Parkinson's Disease.

VI. Conclusion

In conclusion, our investigation into "Parkinson's Disease Detection and Analysis Using Machine Learning" culminates in compelling findings with significant implications for healthcare. The adoption of the Light Gradient Boosting Machine (LightGBM) model has yielded remarkable accuracy, achieving a notable 99.24%. The meticulously curated dataset from the UCI Machine Learning Repository underwent comprehensive analysis and visualization, unraveling essential insights into temporal dynamics within EEG recordings. The model's transparency, as highlighted through feature importance analysis, establishes a crucial link between machine learning outcomes and clinical interpretation. This blend of robust accuracy and interpretability positions our model as a promising tool for practical applications in Parkinson's Disease detection. The seamless integration of advanced machine learning techniques with insightful data analysis not only underscores the effectiveness of our approach but also opens new avenues for transformative solutions in healthcare, particularly in the nuanced realm of Parkinson's Disease diagnosis and monitoring.

Our study not only contributes a valuable model for Parkinson's Disease detection but also signifies the potential of interdisciplinary collaboration between machine learning and healthcare. The synergistic integration of advanced computational methodologies with clinical insights holds promise for revolutionizing the landscape of Parkinson's Disease diagnosis, emphasizing precision and efficiency in patient care. As we advance into an era of innovative healthcare solutions, our findings contribute to the growing body of knowledge bridging the gap between machine learning advancements and tangible benefits for individuals affected by Parkinson's Disease.

References

[1] Luis Alberto Bravo-Vázquez, Erick Octavio Mora-Hernández, Alma L. Rodríguez, Padmavati Sahare, Anindya Bandyopadhyay, Asim K. Duttaroy, Sujay Paul, Current Advances of Plant-Based Vaccines for Neurodegenerative Diseases, Pharmaceutics, 10.3390/pharmaceutics15020711, 15, 2, (711), (2023)

[2] Alnaaim, S.A., Al-Kuraishy, H.M., Alexiou, A. et al. Role of Brain Liver X Receptor in Parkinson's Disease: Hidden Treasure and Emerging Opportunities. Mol Neurobiol (2023). https://doi.org/10.1007/s12035-023-03561-y

[3] Ayaz, Z., Naz, S., Khan, N.H. et al. Automated methods for diagnosis of Parkinson's disease and predicting severity level. Neural Comput & Applic 35, 14499–14534 (2023). https://doi.org/10.1007/s00521-021-06626-y

[4] Haider, A., Elghazawy, N.H., Dawoud, A. et al. Translational molecular imaging and drug development in Parkinson's disease. Mol Neurodegeneration 18, 11 (2023). https://doi.org/10.1186/s13024-023-00600-z

[5] Batiha, G.ES., Al-kuraishy, H.M., Al-Gareeb, A.I. et al. SIRT1 pathway in Parkinson's disease: a faraway snapshot but so close. Inflammopharmacol 31, 37–56 (2023). https://doi.org/10.1007/s10787-022-01125-5

[6] Ali, L., Chakraborty, C., He, Z. et al. A novel sample and feature dependent ensemble approach for Parkinson's disease detection. Neural Comput & Applic 35, 15997–16010 (2023). https://doi.org/10.1007/s00521-022-07046-2

[7] Xu, Y., Li, Y., Wang, C. et al. The reciprocal interactions between microglia and T cells in Parkinson's disease: a double-edged sword. J Neuroinflammation 20, 33 (2023). https://doi.org/10.1186/s12974-023-02723-y

[8] Luis Alberto Bravo-Vázquez, Erick Octavio Mora-Hernández, Alma L. Rodríguez, Padmavati Sahare, Anindya Bandyopadhyay, Asim K. Duttaroy, Sujay Paul, Current Advances of Plant-Based Vaccines for Neurodegenerative Diseases, Pharmaceutics, 10.3390/pharmaceutics15020711, 15, 2, (711), (2023)

[9] Zhu, B., Yin, D., Zhao, H. et al. The immunology of Parkinson's disease. Semin Immunopathol 44, 659–672 (2022). https://doi.org/10.1007/s00281-022-00947-3

[10] Spinelli JB, Haigis MC. The multifaceted contributions of mitochondria to cellular metabolism. Nat Cell Biol20: 745–754, 2018. doi:10.1038/s41556-018-0124-1

[11] C. Bhatt, S. Singh, R. Chauhan, T. Singh and A. Uniyal, "Artificial Intelligence in Current Education: Roles, Applications & Challenges," 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2023, pp. 241-244, doi: 10.1109/ICPCSN58827.2023.00045

[12] TamaraAndrea Alarcón, SarahMartins Presti-Silva, AnaPaula Toniato Simões, FabiolaMara Ribeiro, RitaGomes Wanderley Pires, Molecular mechanisms underlying the neuroprotection of environmental enrichment in Parkinson's disease, Neural Regeneration Research, 10.4103/1673-5374.360264, 18, 7, (1450), (2023)

[13] M. Bajaj, P. Rawat, C. Bhatt, R. Chauhan and T. Singh, "Heart Disease Prediction using Ensemble ML," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 680-685, doi: 10.1109/ICSCDS56580.2023.10104770

[14] Willis, A.W., Roberts, E., Beck, J.C. et al. Incidence of Parkinson disease in North America. npj Parkinsons Dis. 8, 170 (2022). https://doi.org/10.1038/s41531-022-00410-y

[15] Heidari, A., Yazdanpanah, N. & Rezaei, N. The role of Toll-like receptors and neuroinflammation in Parkinson's disease. J Neuroinflammation 19, 135 (2022). https://doi.org/10.1186/s12974-022-02496-w

[16] Espay AJ, Lang AE (2018) Parkinson diseases in the 2020s and beyond: replacing clinico-pathologic convergence with systems biology divergence. J Parkinsons Dis 8(s1):S59–S64. https://doi.org/10.3233/JPD-181465

[17] McFarthing, Kevina | Rafaloff, Garyb | Baptista, Marcoc | Mursaleen, Leahd | Fuest, Rosied | Wyse, Richard K.d | Stott, Simon R.W.d; https://dx.doi.org/10.3233/JPD-229002

[18] Mestre, Tiago A.a; b; * | Fereshtehnejad, Seyed-Mohammadb | Berg, Danielac | Bohnen, Nicolaas I.d | Dujardin, Kathye | Erro, Robertof | Espay, Alberto J.g | Halliday, Glendah | van Hilten, Jacobus J.i | Hu, Michele T.j | https://dx.doi.org/10.3233/JPD-202472

[19] Fanning, S., Selkoe, D. & Dettmer, U. Parkinson's disease: proteinopathy or lipidopathy?. npj Parkinsons Dis. 6, 3 (2020). https://doi.org/10.1038/s41531-019-0103-7

[20] Yuwen Zhao, Lixia Qin, Hongxu Pan, Zhenhua Liu, Li Jiang, Yan He, Qian Zeng, Xun Zhou, Xiaoxia Zhou, Yangjie Zhou https://doi.org/10.1093/brain/awaa167

[21] Jin-Sun Jun, Ryul Kim, Peripheral blood inflammatory cytokines in prodromal and overt α-synucleinopathies: a review of current evidence, encephalitis, 10.47936/encephalitis.2023.00031, 3, 3, (81-86), (2023)

[22] Andrea Ticinesi, Antonio Nouvenne, Nicoletta Cerundolo, Alberto Parise, Tiziana Meschi, Accounting Gut Microbiota as the Mediator of Beneficial Effects of Dietary (Poly)phenols on Skeletal Muscle in Aging, Nutrients, 10.3390/nu15102367, 15, 10, (2367), (2023)