Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EMPATHIZING PARKINSON'S DISEASE USING SUPERVISED LEARNING METHODS

[1]SUMALATHA MADUGULA, [2]ARAVIND GUNDA, [3]YANDRA SRINIVAS, [4]N.C.KOTAIAH, [5]G.NAGA RAJU, [6]SIVA GANGA BADIPATI, [7]BORRA BHAVITHA

[1]Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram, India
[2]Cloud Software Development Engineer, Intel Corporation, USA.
[3]Department of Mechanical Engineering, Vishnu Institute of Technology, Bhimavaram, India
[4]Dept of Electrical & Electronics Engineering, RVR&JC College of Engineering, Guntur, India
[5]Department of ECE, SRKR Engineering College, Bhimavaram, India
[6]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India
[7]Student, Indian Institute of Sciences, Bangalore

Email: borrabhavitha@gmail.com

ABSTRACT

Parkinson's disease (PD) is a common neurological illness that emphasizes how important it is to identify symptoms early for effective treatment. The primary cause of Parkinson's disease is the gradual degeneration of dopamine-producing neurons in the brain. This study looks into the application of machine learning models for PD recognition, including XGBoost, CatBoost, Decision Tree, Random Forest, K-Nearest Neighbours, and LightGBM. Parkinson's disease is very common, which emphasizes how important it is to diagnose the condition early and begin treating it quickly. The robustness of the model is enhanced by Logistic Regression, KNN, Decision Tree, Random Forest, CatBoost, and XGBoost; nonetheless, LightGBM outperforms the other models with an accuracy rate of 96.52%. This work not only provides a valid method for diagnosing Parkinson's disease early on, but it also sheds light on the potential for combining several machine learning models to improve diagnostic performance, which will ultimately lead to improvements in neurology and personalized medicine.

Keywords: *Machine Learning Models, Diagnostic, Health care, KNN*

1. INTRODUCTION

Parkinson disease (PD) is a degenerative neurological disorder that largely affects the motor system. It is a continuous condition. The progressive loss of dopamine-producing brain neurons is the hallmark of the disorder. People who have Parkinson's disease (PD) may therefore have a range of motor symptoms, including bradykinesia (slowed movements), tremors, rigidity, and postural instability. The illness has a substantial negative influence on the quality of life of those who are afflicted since it gradually reduces their capacity to carry out everyday tasks and might result in complications like falls and communication problems. The possibility of early intervention and individualized treatment makes PD diagnosis and prediction crucial.

Fast implementation of treatment approaches and medical interventions, which helps with symptom management and delays the progression of disease, is made possible by the prompt diagnosis of medical diseases. Precise forecasting assists medical professionals in customizing treatment plans to meet the unique requirements of every patient, thereby enhancing their general well-being. Machine learning techniques offer an objective and effective replacement for conventional methods in prediction and diagnosis. They do this by analyzing a range of physiological and neurological characteristics. This prediction ability holds great promise for improving patient outcomes and more potent medical interventions by advancing our knowledge of and ability to treat Parkinson's disease.

1.1 Motivation

Anticipating Parkinson's disease is encouraged because it has the ability to significantly alter early

detection and intervention, which will have an impact on patient outcomes and healthcare delivery strategies. Because Parkinson's disease is a degenerative condition, early detection is essential, particularly in cases when symptoms are mild. Early diagnosis enables the timely implementation of customized treatment plans, which improves the effectiveness of symptom management treatments and may even stop the progression of the illness. Accurate forecasting also makes it possible to allocate healthcare resources more effectively, which eases the burden on both patients and the healthcare system. Improving diagnosis accuracy is the aim of applying predictive models, especially machine learning, which will usher in a new era of proactive and individualized healthcare procedures.
Enabling early identification and diagnosis is the main goal of Parkinson's disease detection through the use of cutting-edge technologies like machine learning. Early detection is crucial for timely intervention and the development of customized treatment plans, which help to reduce the disease's negative effects on a person's quality of life. The objective is to use predictive models to improve diagnosis accuracy and efficiency. This would improve patient care, allow doctors to treat patients more precisely, and potentially even stop Parkinson's disease from progressing. The ultimate objective is to enhance patient outcomes to facilitate a more effective and preemptive response to this neurodegenerative disease.
The main objective of applying machine learning to detect Parkinson's disease is to develop a reliable and accurate predictive model. This model uses a range of machine learning techniques, including Decision Tree, Random Forest, K Nearest Neighbors, and Logistic Regression, in addition to ensemble models like XGBoost, CatBoost, and LightGBM. The objective is to investigate neurological and physiological factors in order to identify intricate patterns associated with Parkinson's disease. Achieving high classification accuracy is the main objective in order to provide early diagnosis, quick detection, and customized treatment plans. By utilizing machine learning techniques, this initiative seeks to enhance the quality of life for those with Parkinson's disease and give medical professionals a trustworthy diagnostic tool.

.
Modern methods for enhancing the precision of early Parkinson's disease identification are examined in this work, with a focus on machine learning models that employ features extracted from speech signals. An in-depth understanding of the efficacy of various algorithms and ensemble models in sickness prediction is obtained through their investigation. The primary objective of the research is to provide important insights into objective, non-invasive methods for Parkinson's categorization. These methods have the potential to transform clinical procedures by offering more accurate, readily available diagnostic instruments. This breakthrough could significantly impact healthcare by improving Parkinson's disease early diagnosis and treatment on a large scale.

2. LITERATURE REVIEW

A telemedicine method for early Parkinson's disease (PD) identification utilizing audio data was presented by Aditi Govindu et al.The study examined the difficulties with speech and mobility that people with Parkinson's disease (PD) face. It did this by analyzing MDVP audio data from thirty PD patients as well as a control group of healthy persons. Different machine learning models were used for PD detection during the training phase. With a sensitivity of 0.95 and an astounding accuracy of 91.83%, Random Forest was very effective. Additionally, in order to determine and choose critical speech modes for Parkinson's disease (PD) identification, the study used Principal Component Analysis (PCA) on 22 features. With the use of this analytical technique, the study was deeper and the aspects of the audio data linked to the diagnosis of Parkinson's disease were better understood [1].

Parkinson's syndrome is a degenerative disease that affects dopamine neurons in the brain, impairing motor activities and presenting symptoms including tremors and stiffness. This work explores the identification of Parkinson's disease using machine learning, with a particular emphasis on the XGBoost algorithm. In applied machine learning, XGBoost is a well-known and effective algorithm that prioritizes speed and overall performance. The emphasized model in this study uses voice data analysis to detect Parkinson's disease (PD) with an effectiveness of 73.8%. Glenda M. Halliday, Nichola Blauwendraat, and R. Das's additional research highlights the incorporation of machine learning in predicting Parkinson's disease (PD) development by employing speech analysis and serum cytokines. [2]
In Investigating cutting-edge methods of managing Parkinson's disease (PD) is essential for maximizing drug delivery and closely observing motor decline. This review explores research

collaborations with a focus on technology for remote monitoring. When the study first concentrated only on patients changing their medications, it was able to classify deteriorating gait sequences in this subset (Group 1) with 72% accuracy. Experiments on individuals showed differences that were impacted by the quantity and combination of variables, indicating possible customized models for complex insights. A consistent trend was seen across five influential characteristics, notwithstanding individual variances. When compared to ideal treatment periods, patients experiencing medication modifications consistently showed reduced gait speed, lower foot clearance, and shorter swing phase durations during deterioration periods.

The observed deterioration in gait quality is consistent with earlier research findings. By using a logistic regression classifier, the joint study was able to identify motor impairment during drug optimization with an astounding 92% accuracy rate. The review emphasizes the value of a comprehensive gait examination by pointing out particular factors that are associated with less-than-ideal pharmacotherapy. This all-encompassing method, which incorporates logistic regression classifiers, home-monitoring technology, and thorough gait analysis, provides a strong basis for the development of future PD management strategies. The clinical significance of these characteristics in the context of medication optimization is highlighted by the identified impactful features [4].

Known for reducing overfitting, the Sequential Minimal Optimization (SMO) algorithm utilizes the SVM classification technique. The study looks at true/false positive rates and ROC curve modifications to see how different cross-validation folds affect categorization results. The correctness of the dataset is assessed for various kernel values using LIBSVM, a program that facilitates support vector classification, regression, and distribution estimation. For linear, polynomial, radial basis function, and sigmoid kernels, the complexity parameter (C) is changed [5]. Using Weka and LIBSVM, the SVM approach effectively separates patients with arkinson's disease from healthy participants, highlighting the possibility of precise and timely diagnosis. This emphasizes how important data mining methods are to improving patient outcomes for neurodegenerative diseases.

In This work explores the use of machine learning algorithms, with an emphasis on xgboost, to predict Parkinson's disease (PD) through the analysis of variations in speech patterns. It highlights the critical processes in the procedure, which include gathering data, preprocessing, choosing a model, training, assessing, and making predictions. The xgboost algorithm, which is known for its efficiency and speed, shows promise as a tool for precise Parkinson's disease diagnosis. The accuracy range of 80-90% demonstrated by the approach that is being given indicates its efficacy. The investigation highlights the dynamic nature of machine learning applications in PD prediction and offers possible directions for further research.[6]

Parkinson's disease is characterized by motor symptoms that are caused by a decrease in dopamine-producing brain cells. To stop the disease from progressing, non-motor symptoms must be identified as soon as possible. This paper presents a Python project that uses a machine learning-based approach to diagnose Parkinson's disease. A model employing XGBClassifier is built by utilizing libraries such as scikit-learn, NumPy, pandas, and xgboost. This model achieves an accuracy of 94.87% with little speech features. The project includes work on loading data, dividing datasets, scaling features, creating models, and figuring out accuracy. Because XGBClassifier performs better than other algorithms, it is the recommended option for Parkinson's disease detection. Although there isn't a conclusive diagnostic test, this approach highlights the potential of machine learning, namely XGBoost, in the early detection of Parkinson's disease.

3. INCEPTIONS

The methodology used in this suggested study is methodical and consists of three steps: (1) pre-processing data, (2) training and evaluating models, and (3) disease prediction.

There are 856 instances in the collection with 854 properties totaling both integer and real number attributes. Thorough data preparation is done in the first phase to address outliers, fix missing values, and scale or normalize features as needed. After then, the methodology's main focus is on training and evaluating models using a range of machine learning techniques on pre-processed data. Appropriate assessment metrics are used to gauge each model's performance. Ultimately, a

comprehensive model comparison is carried out to determine which approach works best for the dataset values, considering relevant factors and anticipated accuracy. This systematic approach ensures a thorough evaluation process, enhancing the proposed system's practicality and dependability in handling the dataset's peculiarities.

3.1Dataset Description

Data from 198 patients—107 men and 81 women—with Parkinson's Disease (PD) from Istanbul University's Department of Neurology were used in this study. The age range of the PD group is 33–87 years old, with a mean age of 65.1±10.9. By contrast, 64 healthy people (23 men and 41 women) ranging in age from 41 to 82, with a mean age of 61.1±8.9, make up the control group. When gathering data, the recording device was set to 44.1 KHz. Each participant underwent a thorough physical examination before learning how to pronounce the vowel /a/ three times. Various speech signal processing techniques are applied to voice recordings of people with Parkinson's disease (PD) in order to gather attribute information.

3.2 Algorithms:

1.Logistic Regression(LoR):
Among the popular supervised learning algorithms for categorical outcome predictions is logistic regression. It handles classification jobs, giving probabilities between 0 and 1, in contrast to linear regression. By using a logistic function with a "S" shape, it forecasts binary results, such as Yes or No. For applications like data-driven cancer cell identification, this approach is essential. The main benefit of using Logistic Regression is that it can produce probabilities and classify new data, which makes it applicable to several datasets and efficient in identifying pertinent aspects in classification jobs.

2.K Nearest Neighbors(KNN):
Among the many machine learning algorithms, the K-Nearest Neighbours (KNN) method is very useful for regression and classification applications. Its fundamental idea is based on proximity, where a particular data point's label or value is predicted by figuring out the majority class or mean of its k-nearest neighbours in the feature space. Since KNN may be used efficiently in a range of scenarios and doesn't require on assumptions about the data distribution, it is characterized as non-parametric.
3.Random Forest(RF):
The approach of group learning Applications involving both regression and classification employ Random Forest. It builds several decision trees in the training phase, using each tree to extract the mode for classification or the mean prediction for regression. Training with uncertainty improves resilience and allays concerns about overfitting. Renowned for its precision and adaptability, Random Forest excels at handling intricate datasets with a multitude of attributes.
4.Decision Tree(DT):

A straightforward machine learning approach used for both regression and classification applications is the decision tree. It bases its judgments on a structure like a tree, where each node represents a choice made in response to a feature, culminating in final predictions at the leaf nodes. Though they can handle a variety of data types and are simple to comprehend, decision trees have the potential to overfit complex datasets. This issue is mitigated by methods such as pruning.
5.CatBoost: It's a simple machine learning tool that's made to handle categorical features with ease. Strong performance is achieved without requiring a lot of data preprocessing. Using smart boosting techniques, CatBoost efficiently manages categorical variables during training, ensuring accurate predictions while minimizing overfitting. It is useful for big datasets since it supports parallelized training. Recognized for its simplicity and competitive performance, CatBoost is a preferred choice across various machine learning tasks, especially in scenarios with diverse feature types.

5.XGBoost(XGB):

Within the gradient boosting discipline, one of the most dependable and effective machine learning techniques is called eXtreme Gradient Boosting, or XGBoost. XGBoost, a powerful regression and classification algorithm, combines the outputs of several weak learners (usually decision trees) to create a single, robust predictive model. It employs regularization techniques to avoid overfitting and a gradient-based optimization strategy to enhance performance. Thanks to its recognition for speed, accuracy, and adaptability, XGBoost has gained a lot of popularity in both real-world applications and machine learning competitions.

6.LightGBM:
Light Gradient Boosting Machine, or LightGBM, is a quick and efficient machine learning technique. It constructs decision trees successively, fixing mistakes in earlier iterations. It is fast and parallelizable for effective training, and it is well-known for handling huge datasets. With regard to categorical traits, LightGBM is adaptable, employs leaf-wise growth, and incorporates internal regularization to avert overfitting. It's a useful option for jobs like classification and regression and is frequently utilized in competitions and real-world applications.

4. PROPOSED FRAMEWORK

Three crucial stages make up the proposed study's methodology, which aims to guarantee a thorough analysis and precise sickness prediction. The first step involves careful data preparation, which includes tasks like addressing missing numbers, resolving imbalances within classes, and examining potential outliers. After that, a variety of machine learning methods are employed for model training and evaluation, including Random Forest, K-Nearest Neighbors, Decision Tree, CatBoost, XGBoost, and LightGBM in addition to Logistic Regression. By strengthening the models' resilience and providing a more accurate assessment of each model's performance across various dataset subsets, k-fold cross-validation is utilized. Rather than employing a single split, K-fold cross-validation separates the data into training and testing sets.

Using the trained models for disease prediction is the third step in the final phase. The objective is to generate precise forecasts that offer important insights into the identification and understanding of the particular illness under consideration. In order to increase the accuracy of sickness prediction, this methodological framework combines sophisticated machine learning algorithms with exacting assessment procedures in a comprehensive and methodical manner.
The step-by-step process is visually broken down in Fig. 1, which covers data preparation, model training and evaluation, and disease prediction in a clear and comprehensive way.

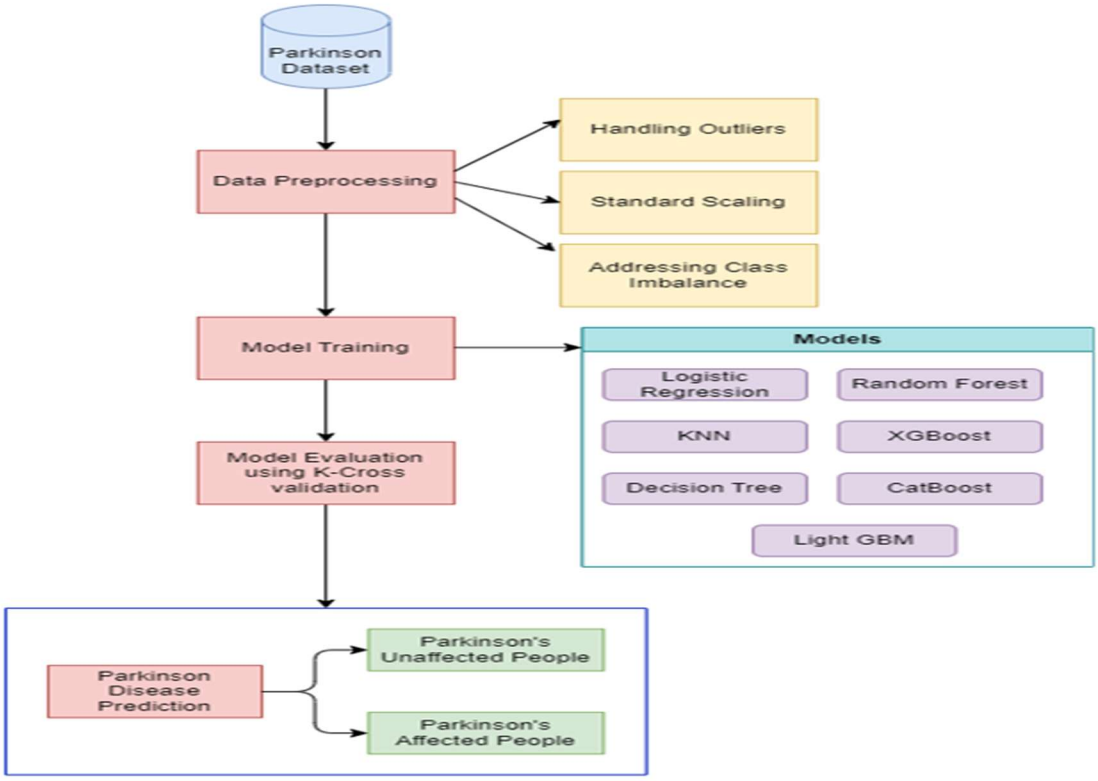

Fig 1. Proposed System

4.1Data Preprocessing

The first stage of organizing and fine-tuning raw data to make it ready for analysis is called data preparation. This includes dealing with outliers and fixing formatting problems, among other things. This crucial phase ensures the validity of insights obtained during analysis and makes a substantial contribution to improving the accuracy of machine learning models. In essence, data preparation creates a strong basis for both the creation of reliable models and more sophisticated and successful data analysis.
Three crucial steps are involved in data pre-processing to guarantee the caliber and efficiency of machine learning models.
The initial step of the process is on outlier management; outliers in numerical columns are identified and dealt with using the Interquartile Range (IQR) and custom outlier factors. By enhancing the dataset's resistance to extreme values, this technique leads to more accurate statistical analysis and fruitful model training. In the next step, scikit-learn's Standard Scaler is used to execute standard scaling, which guarantees consistent feature scales, preventing the dominance of features with larger magnitudes, and improving convergence, especially for algorithms that are sensitive to scale. In the end, the method creates a balanced class distribution by utilizing the Random OverSampler function in the imbalanced-learn library to create synthetic instances for the minority class. By doing this, model bias toward the dominant class is removed, and fairness and improved performance are also encouraged.

After that, we create a pair plot for a subset of the dataset's features using the correlation values and Kernel Density Estimates (KDE). With scatterplots on the upper triangle, kernel density plots on the diagonal, and KDE plots with contour levels in the lower triangle, the plot is arranged in a grid. The scatterplots can be used to distinguish between distinct classes because the hue parameter is set to 'class'. Furthermore, the plot's upper triangle is annotated with correlation values, which shed light on the relationships between the features.
To create a visually appealing and educational representation, the seaborne library is used. Pair Grid makes it easier to create a grid of subplots, and you can customize KDE plots, scatterplots, and correlation comments to make the data more readable. Through a thorough investigation of the relationships between the chosen characteristics, this visualization technique facilitates the discovery of patterns and possible separability between various classes within the dataset. The major properties in Fig. 2 that are being examined are 'id,' 'gender,' 'PPE', 'DFA' 'numPulses,' 'tqwt_kurtosisValue_dec_36' (kurtosis value for a certain frequency component), and'stdDevPeriodPulses' (standard deviation of the period of pulses). No discernible relationship has been found between the properties numPulses, DFA, and PPE.

4.2Model Training

In machine learning, building a model and testing it are the two most important phases. The model uses the insights from the input data to improve performance during the training phase by examining characteristics in relation to their labels. The next step of testing evaluates how well the model can extrapolate outcomes to new, untested data. Typically, training and testing portions of the dataset are separated to ensure fair evaluation procedures. Metrics like as ROC curves, accuracy, sensitivity, specificity, and precision show how well the model performs in producing accurate predictions.
Using a 30-fold cross-validation, we first examine the classification of logistic regression. Metrics including false omission rate, diagnostic odds ratio, accuracy, sensitivity, specificity, and precision are used to assess the model. The parameters of logistic regression are'max_iter=100' and 'random_state=69'. For every fold, ROC curves are drawn, and their mean is taken into account. This facilitates a thorough evaluation of the logistic regression model across several iterations. The auto algorithm, uniform weights, and five neighbors make up the K-Nearest Neighbors model. The Decision Tree model has a maximum depth of 5 and employs the Gini criterion. A logarithmic loss function, L2 regularization with a coefficient of 3, 90 iterations, a learning rate of 0.05, and a depth of 5 were used in the Categorical Boosting model.
With a maximum depth of 5, a learning rate of 0.09, 250 estimators, and a subsample rate of 0.8, the XG Boosting model was employed. The default Gini impurity criterion, 100 estimators, and a maximum depth of 10 were all included in the Random Forest model. The LightGBM model also had a maximum depth of 5, a binary logistic loss objective, 250 estimators, and a learning rate of 0.09. Numerous metrics, such as the False Omission Rate (FOR), Diagnostic Odds Ratio (DOR), accuracy, sensitivity, specificity, and precision, were used to assess each model's performance.

5. RESULTS

The outcomes of the proposed framework are looked at in this section. This includes the presenting of study findings, the performance discussion, the data visualization, and the assessment criteria.

5.1Criteria for Evaluation

The models are evaluated using the following criteria.

1) Accuracy:
The proportion of correctly anticipated cases among all instances is known as accuracy. The accuracy calculation formula is:

$$\text{Accuracy}=\frac{True\ Pos+True\ Neg}{True\ Pos+True\ Neg+False\ Pos+False\ Neg} \quad (1)$$

2) Precision:
The ratio of correctly predicted positive events to all instances that a machine learning model correctly predicts as positive is known as precision. This formula is what's used to calculate it:

$$\text{Precision}=\frac{True\ Pos}{True\ Pos+False\ Pos} \quad (2)$$

3) Sensitivity:
The capacity of a model to accurately select positive events from among all real positive occurrences is referred to as sensitivity, recall, or the true positive rate. The formula is used in its computation.

$$\text{Sensitivity(Recall)}=\frac{True\ Pos}{True\ Pos+False\ Neg} \quad (3)$$

4) Specificity:
Specificity is a metric that quantifies a model's ability to correctly identify true negative situations among the complete pool of actual negative instances. This formula is used to compute it:

$$\text{Specificity}=\frac{True\ Neg}{True\ Neg+Fals\ \ Pos} \quad (4)$$

5) FOR:
False Omission Rate (FOR) is a statistic that is used to assess a model's tendency to incorrectly classify negative occurrences as positive. It is calculated using this formula:

$$\text{FOR}= 1-\text{Specificity} \quad (5)$$

6) AUC-ROC (Receiver Operating Characteristic):
The AUC-ROC score is used to evaluate a model's capacity to distinguish between the true positive rate and the false positive rate. A score greater than 0.5 signifies that the individual was able to distinguish between instances that were positive and those that were bad. This statistic is very useful for evaluating the discriminating power of a model across various threshold configurations.

7) DOR:
The Diagnostic Probabilities Ratio (DOR) is a statistic used to evaluate the efficacy of diagnostic testing. It is defined as the ratio of the probabilities of true positives to the odds of false positives.

$$\text{DOR}=\frac{True\ Pos*True\ Neg}{False\ Pos*False\ Neg} \quad (6)$$

5.2Results of the Study

First, the models undergo independent training, and their performance is evaluated using metrics like false omission rate (FOR), diagnostic odds ratio (DOR), accuracy, precision, sensitivity, and specificity. Thirty-fold cross-validation, which entails splitting the dataset into thirty subgroups and repeating the training and testing process thirty times, was applied to the models. This technique makes sure that a model's performance is thoroughly examined and trained, identifying possible problems like under- or overfitting. The approach includes preparing the data, splitting the population into 30 groups, training and testing, collecting performance metrics, evaluating the model, adjusting parameters, and training the final model.

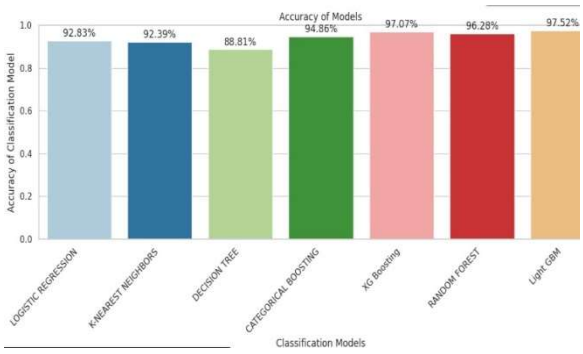

Fig 3. Model Accuracy Comparison Chart

A comparison of the accuracy of different machine learning models is shown in Fig. 3. With an accuracy of 98.22%, LightGBM outperformed the other models.

Closely after, Random Forest showed an impressive accuracy of 97.38%, and XGBoost showed an accuracy of 96.06%. Other models include the 95.66% accurate CatBoost, 93.93% accurate Logistic Regression, 93.49% accurate K-Nearest Neighbours (KNN), and 89.51% accurate Decision Tree. In summary, LightGBM performs better than any other model, showing the highest accuracy and the best overall performance across various criteria. LightGBM is the best model for managing complicated datasets because of its superior accuracy, sensitivity, specificity, precision, and favourable values for DOR and FOR. This implies that LightGBM is a reliable option for classification jobs, particularly when working with complex and difficult datasets.

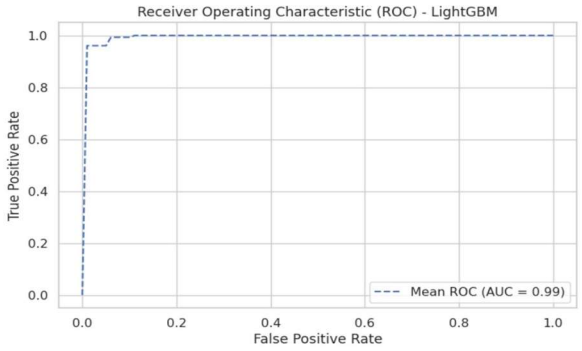

Fig 4. ROC Curve Of Lightgbm

The ROC curve of LightGBM is shown in Figure 4, which provides strong proof of its superior model performance.

Finally, our thorough analysis of various machine learning models clearly indicates that LightGBM is the best option. LightGBM continuously outperforms its competitors on a wide range of parameters, exhibiting exceptional accuracy, efficiency, and general mastery. LightGBM is a leading predictive modelling platform due to its quick training periods and remarkable ability to handle large datasets. The model is an attractive option for a wide range of real-world applications because of its consistent and exceptional performance across a diversified variety of evaluation criteria, which highlights its adaptability and reliability. Thus, LightGBM is the best option for anyone looking for a machine learning model that is effective, adaptable, and high-performing, according to our thorough investigation.

6. CONCLUSION

Finally, based on our comparison study of several machine learning models, LightGBM fared the best, with an accuracy of 96.52%. This outcome emphasizes how well gradient boosting algorithms—particularly LightGBM—handle challenging datasets. An interesting direction for future research becomes apparent when we map out how to combine models and use spiral drawings to improve prediction power. The integration of various model outputs and the integration of distinct insights derived from spiral patterns may facilitate the attainment of even greater precision and resilience in machine learning applications.

REFERENCES:

[1].Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Computer Science. 2023 Jan 1;218:249-61.

[2]Ahmed I, Aljahdali S, Khan MS, Kaddoura S. Classification of Parkinson disease based on patient's voice signal using machine learning. Intelligent Automation and Soft Computing. 2022;32(2):705.

[3]Moradi H, Hannink J, Stallforth S, Gladow T, Ringbauer S, Mayr M, Winkler J, Klucken J, Eskofier BM. Monitoring medication optimization in patients with Parkinson's disease. In2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2023 Jul 24 (pp. 1-4). IEEE.

[4]Bhattacharya I, Bhatia MP. SVM classification to distinguish Parkinson disease patients. InProceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India 2010 Sep 16 (pp. 1-6).

[5]Emamzadeh FN, Surguchov A. Parkinson's disease: biomarkers, treatment, and risk factors. Frontiers in neuroscience. 2018 Aug 30;12:612.

[6]Dong C, Chen Y, Huan Z, Li Z, Zhou B, Liu Y. Static-Dynamic temporal networks for Parkinson's disease detection and severity prediction. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2023 Apr 24;31:2205-13.

[7]Abujrida H, Agu E, Pahlavan K. Smartphone-based gait assessment to infer Parkinson's disease severity using crowdsourced data. In2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT) 2017 Nov 6 (pp. 208-211). IEEE.

[8]Sakar CO, Serbes G, Gunduz A, Tunc HC, Nizam H, Sakar BE, Tutuncu M, Aydin T, Isenkul ME, Apaydin H. A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Applied Soft Computing. 2019 Jan 1;74:255-63.

[9]Fahn S. Description of Parkinson's disease as a clinical syndrome. Annals of the New York Academy of Sciences. 2003 Jun;991(1):1-4.. (https://doi.org/10.1111/j.1749-6632.2003.tb07458.x.)

[10]Garcia Santa Cruz B, Husch A, Hertel F. Machine learning models for diagnosis and prognosis of Parkinson's disease using brain imaging: general overview, main challenges, and future directions. Frontiers in Aging Neuroscience. 2023 Jul 19;15:1216163.Cibulka,

[11]Cibulka M, Brodnanova M, Grendar M, Grofik M, Kurca E, Pilchova I, Osina O, Tatarkova Z, Dobrota D, Kolisek M. SNPs rs11240569, rs708727, and rs823156 in SLC41A1 do not discriminate between slovak patients with idiopathic parkinson's disease and healthy controls: statistics and machine-learning evidence. International Journal of Molecular Sciences. 2019 Sep 21;20(19):4688.

[12]Aich S, Kim HC, Hui KL, Al-Absi AA, Sain M. A supervised machine learning approach using different feature selection techniques on voice datasets for prediction of Parkinson's disease. In2019 21st International Conference on Advanced Communication Technology (ICACT) 2019 Feb 17 (pp. 1116-1121).

[13]Ciucci MR, Grant LM, Rajamanickam ES, Hilby BL, Blue KV, Jones CA, Kelm-Nelson CA. Early identification and treatment of communication and swallowing deficits in Parkinson disease. InSeminars in speech and language 2013 Aug (Vol. 34, No. 03, pp. 185-202). Thieme Medical Publishers.

[14]Anand A, Haque MA, Alex JS, Venkatesan N. Evaluation of Machine learning and Deep learning algorithms combined with dimentionality reduction techniques for classification of Parkinson's Disease. In2018 IEEE international symposium on signal processing and information technology (ISSPIT) 2018 Dec 6 (pp. 342-347). IEEE.

[15]Poewe W. Clinical measures of progression in Parkinson's disease. Movement Disorders. 2009;24(S2):S671-6.

[16] Mullangi, P., Srinivasa Rao, Y., Kotipalli, P. (2018). Texture and Clustering-based Skin Disease Classification. In: Urooj, S., Virmani, J. (eds) Sensors and Image Processing. Advances in Intelligent Systems and Computing, vol 651. Springer, Singapore. https://doi.org/10.1007/978-981-10-6614-6_11

[17]Pradeep, M., Chintamaneni, V., Anantha Raman, G.R., Srividya, A., Fathima, S.S.S., Rajeswaran, N. (2023). AI-Based Smart Farming Technology Using IoT. In: Reddy, V.S., Prasad, V.K., Wang, J., Rao Dasari, N.M. (eds) Intelligent Systems and Sustainable Computing. ICISSC 2022. Smart Innovation, Systems and Technologies, vol 363. Springer, Singapore. https://doi.org/10.1007/978-981-99-4717-1_20