Analysis and Prediction of Parkinson's Disease using Machine Learning Algorithms

Sohom Sen, Student, Department of Computer Science and Engineering, Heritage Institute of Technology, Kolkata, India, Ankit Ghosh, Student, School of Nuclear Studies and Application, Jadavpur University, Kolkata, India

***Abstract*— Even with radical advancement of medical science, there are medical abnormalities in human beings which are still incurable. First-time correct diagnosis of such ailment helps in decelerating the deterioration of a patient. With the advent of Artificial Intelligence (AI) and Machine Learning (ML), nearly accurate diagnosis of various medical abnormalities is achievable; helping ~~the~~ Doctors to a great extent. Parkinson's Disease is one such ailment for which no cure exists till date. It causes the malfunctioning of our Central Nervous System (CNS) leading to Tremor in limbs, Slowed Physical Movement, Rigid Muscles, Impaired Posture and Balance, Loss of Automated Movements and sometimes Loss of Speech. This article focuses on implementing State-Of-The-Art Machine Learning Models namely, K-Nearest Neighbour (KNN), Logistic regression (LR), Random Forest Classifier (RFC) and XGBoost Classifier (XGBC) for near accurate diagnosing of whether a patient suffers from Parkinson's Disease or not. The ML Models are implemented on Oxford Parkinson's Disease Detection Dataset. The following metrics were used for the performance analysis of implemented ML Algorithms: Precision, Recall, F1-Score, Accuracy, AUC-ROC Curve and AUC-PR Curve. The Performance Analysis revealed, XGBoost Classifier to have an overall highest Accuracy of 96.7%, Precision of 0.8571, Recall of 1.00 and F1-Score of 0.923.**

***Index Terms*— Artificial Intelligence, Binary Classification Model, Extra Trees Classifier, Feature Selection, Machine Learning, Mutual Information, Parkinson's Disease, Performance Analysis**

I. Introduction

Parkinson's Disease is a genre of neurodegenerative disorder affecting the Central Nervous System (CNS) and dopamine-generating neurons in Human Brain. Till date, there has been no cure discovered for this particular ailment. A person suffering from Parkinson's will have to live with it till the end of his/her life. With the development of Medical Science, we have been able to achieve how to decelerate the deterioration of the patient. The reason for the development of such an incurable ailment in a person is still not known to us. Thus, the best solution available for us, till date, is first-time correct diagnosis of Parkinson's Disease.

Deployment of Artificial Intelligence (AI) and Machine Learning (ML) in the field of Medical Science have helped billions of people. The rise of Automated Detection not only assists in near accurate diagnosis of an ailment, but also classifies them into different categories for the ease of medication.

In this article, we have deployed four (04) Machine Learning Models namely, K-Nearest Neighbour (KNN), Logistic Regression (LR), Random Forest Classifier (RFC) and XGBoost Classifier (XGBC) to classify whether a Patient is suffering from Parkinson's Disease or not (Binary Classification Model). All the implemented models went through a Performance Analysis Test based upon distinguished metrics.

II. Related Work

Raundale et al. [1] presented the use of Machine Learning and Deep Learning Algorithm to predict the severity of Parkinson's Disease. They have used UCI's Parkinson's Telemonitoring Vocal Data Set of patients in implementing their proposed algorithms.

Maunika and Rao [2] proposed the use of K-Nearest Neighbour (KNN), State-Of-The-Art Supervised ML Algorithm with the value of k=5 to achieve an accuracy of 97.43%. They have also implemented Deep Learning Algorithms to obtain substantial results in predicting the severity of Parkinson's Disease.

Ranjan and Swetapadma [3] implemented Support Vector Machine (SVM), K-Nearest Neighbour (KNN) and Artificial Neural Network (ANN) for the detection of Parkinson's Disease. After carrying out a comparative study among these three proposed methodology it was found to have an accuracy of around 100% for the testing dataset.

Conventionally, monitoring of the severity of Parkinson's Disease is conducted by Physicians or Medical Representatives at medical centres. Chén et al. [4] proposed a new path for remotely assessing Parkinson's Disease using Smartphones. Postural instability, Dexterity, Gait, Tremor and Voice are recorded by the smartphone continuously. A two-step feature selection process is incorporated inside the implemented framework. Utilizing all these data, they have been able to achieve substantially promising results.

Zhang et al. [5] opted for a two-step method which are Statistical Analysis and Machine Learning method for extracting the differential regions of the brain from the structural MRI data. They were able to achieve the highest accuracy of 93.75% in classifying the severity of Parkinson's Disease.

The use of two Convolution Neural Network (CNN) models, Inception V3 and ResNet50 was implemented along with transfer learning method to detect Parkinson's Disease. This methodology was proposed by Jahan et al. [6]. They were able to achieve an accuracy of 96.67% on the Inception V3 model.

Bhan et al. [7] aimed in diagnosing Parkinson's Disease using Image Enhancement, Region Of Interest Extraction and Deep Learning Algorithms. Substantia Nigra (SN) and VGG-16 CNN algorithms were implemented on MRI Scans to extract pivotal features. ResNet-34, VGG-19 and ResNet-50 Deep Learning Algorithms were implemented on the extracted features to classify Parkinson's Disease. After a comparative analytical study, they concluded, ResNet-50 to have the highest accuracy.

With the use of Convolution Neural Network (CNN) and Artificial Neural Network (ANN), Ouhmida et al. [8] on the UCI dataset with 45 acoustic features were able to achieve an highest accuracy of 93.10% . They have tested their model on two separate repository databases, labelled as Database I and Database II.

III. Dataset

Oxford Parkinson's Disease [9] Dataset have been used to showcase the work. There are a total of 1195 entries in the dataset, each having 23 attributes. The dataset consists of 901 patients correctly diagnosed with Parkinson's Disease and 295 remaining are healthy people, not suffering from the ailment.

For the sake of better understanding, a brief description of each attribute is listed as follows :

1) MDVP:Fo(Hz)
It is the average vocal fundamental frequency of the patients.

2) MDVP:Fhi(Hz)
It is the maximum fundamental vocal frequency of the patient.

3) MDVP:Flo(Hz)
It is the minimum fundamental vocal frequency of the patient.

4) MDVP:Jitter(%), MDVP:Jitter(Abs), MDVP:RAP, MDVP:PPQ, Jitter:DDP
They are the various measures of fluctuations in the fundamental vocal frequency of the patient.

5) MDVP:Shimmer, MDVP:Shimmer(dB), Shimmer:APQ3, Shimmer:APQ5, Shimmer:DDA
They are the various measures of fluctuation on the amplitude of the fundamental frequency.

6) NHR, HNR
These are defined as the measures of the ration of noise to tonal portion in the voice of the patient.

7) RPDE, D2
These are defined as the non-linear complexity measures.

8) DFA
It is defined as the signal fractal scaling component.

9) Spread1, Spread2, PPE
They are defined as the three non-linear measures of the fundamental frequency variation of the patient.

10) Status
The patients suffering from Parkinson's Disease are labelled as 1 and the remaining ones are labeled as 0.

In spite of using a relatively small dataset, the implemented Machine Learning Models [10] came up with impressive results. With inundation of more data, our implemented Models are likely to improve even further.

IV. State-Of-The-Art Machine Learning Algorithms

The Machine Learning Models implemented belongs to the category of Supervised Learning Methods. In this method, we train our Machine Learning Models using labeled input and output data. The Model then further adjusts its weights and biases according to the best fit, to yield the highest accuracy feasible.

A brief description of the implemented State-Of-The-Art Machine Learning Algorithms:

1) K-Nearest Neighbour (KNN)
The algorithm is most commonly used for Classification Models. After the training of the algorithm, it completes the classification of the data into desired number of groups or classes. The testing element is picked out from the pile and compared with all the neighboring elements present. The group with which the testing element has the highest resemblance is declared as the output. This comparison process is basically a voting mechanism. Fig. 1 illustrates a simple picture for the working of KNN.

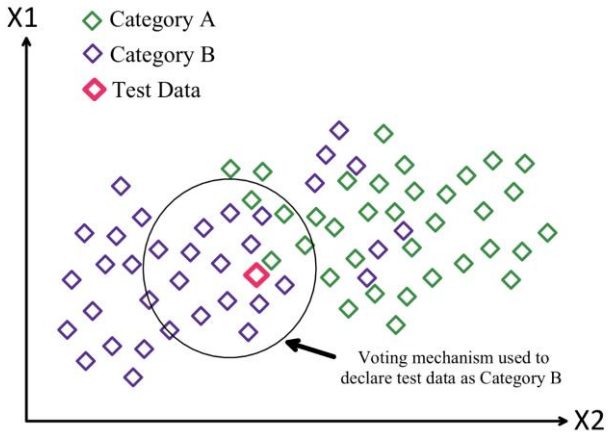

Fig. 1. K-Nearest Neighbour Working Principle

2) Logistic Regression (LR)
The algorithm is primarily used for Binary Classification Models, despite carrying the name regression. The algorithm utilizes a Logistic Function defined as $1/(1+e^{-x})$ and loss function coined as Maximum Likelihood Estimation (MLE). The probability of the input data is computed while training the model. When the test data has a probability greater than or equal to 0.5, it is labelled in Class A otherwise in Class B for example. Brief illustration is shown in Fig.2 for better visualization.

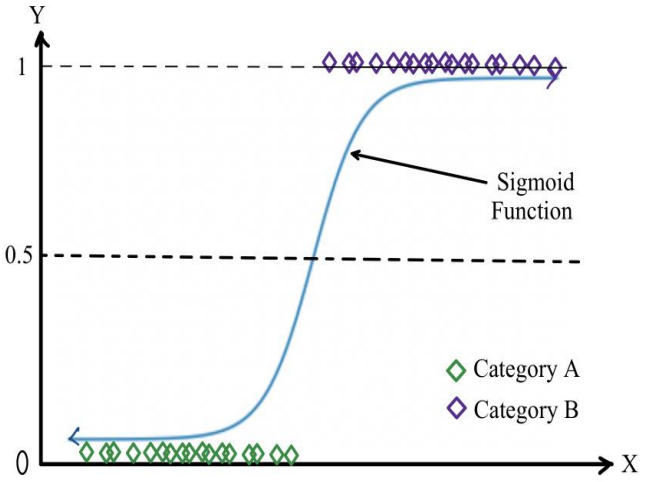

Fig. 2. Logistic Regression Working Principle

3) Random Forest Classifier (RFC)

Decision Trees are the building blocks of this algorithm. Each node of the tree helps in creating distinguished classes to categorize our input data. Random Forest Classifier creates multiple Decision Tree Models with different weights and biases to train them simultaneously and then considers the model with the highest accuracy. This method of simultaneous training of similar models under the roof of an algorithm is called the Bagging Technique. The process is illustrated in Fig. 3.

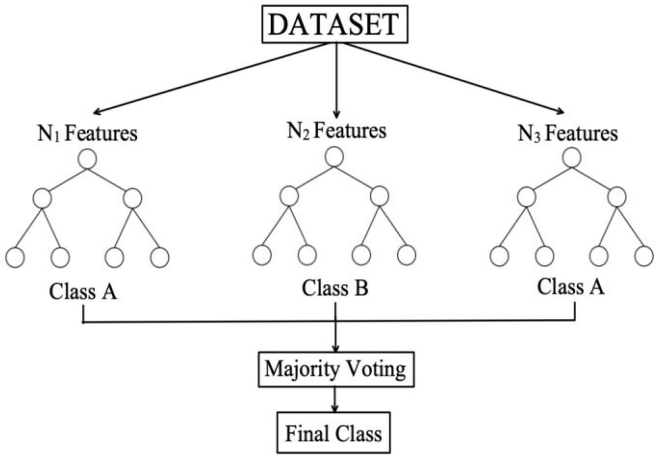

Fig. 3. Random Forest Classifier Working Principle

4) XGBoost Classifier (XGBC)

This Algorithm follows the Boosting Technique. Small sub-models are created with the labelled input data. The models are then trained with specific weights and biases. The model yielding low accuracy are reconsidered with a new set of weights and biases. The process continues in a loop till the models yield a substantial accuracy. Thus, the models are 'boosted'. The Boosting Process is illustrated in Fig. 4.

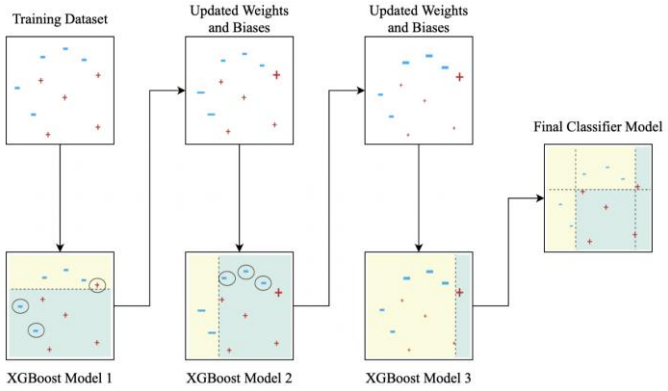

Fig. 4. Boosting Process for XGBoost Classifier

V. Methodology

All the steps involved in predicting whether a person is suffering from Parkinson's Disease or not is described below ~~is~~ as illustrated in Fig. 5.

Flowchart Of Methodology Proposed

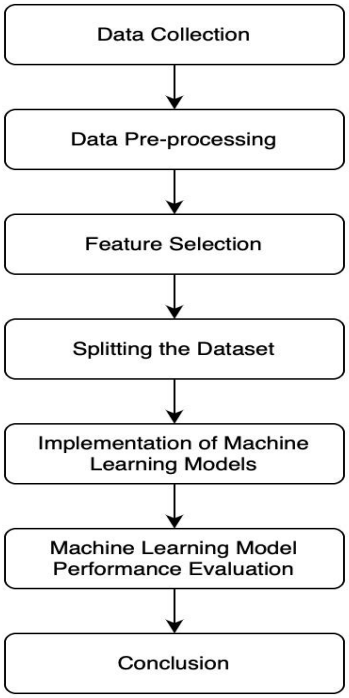

Fig. 5. Flowchart of Methodology proposed

A. Data Collection

The data is obtained from Oxford Parkinson's Disease Dataset.

B. Data Pre-processing

After importing the dataset, [11], [12] the first thing to be checked is presence of Null Values (NaN). To deal with the problem, either the entries containing NaN values are removed from the dataset or the NaN values are replaced by the

mean/median of that particular column if feasible. The second stated method is applied in the dataset.

The following step includes checking the skewness of attributes (how badly the Gaussian Distribution [13] of each attribute is distorted from its mean position). The Skewness [14] of the attributes are checked mainly in two ways. The first one mainly deals with the distribution plot of each attribute. The second and the more reliable way is to manually calculate the skewness of each attribute. Depending on the distorted distribution and the skewness value obtained, it can either be Left-Skewed or Right-Skewed. The skewness [15] can be reduced by performing a combination of the following functions:

- Logarithmic Function
- Square Function
- Square-root Function

The attributes possessing skewness value greater 1.5 or less than -1.5 have been portrayed in Table I.

TABLE I
ATTRIBUTES WITH HIGH SKEWNESS VALUE

ATTRIBUTES	SKEWNESS
MDVP:Fhi(Hz)	2.542
MDVP:Jitter(%)	3.085
MDVP:Jitter(Abs)	2.649
MDVP:RAP	3.361
MDVP:PPQ	3.074
Jitter:DDP	3.362
MDVP:Shimmer	1.665
MDVP:Shimmer(dB)	1.999
Shimmer:APQ3	1.581
Shimmer:APQ5	1.799
MDVP:APQ	2.618
Shimmer:DDA	1.581
NHR	4.221

The Distribution Plot for selected skewed attributes are illustrated in Fig. 6.

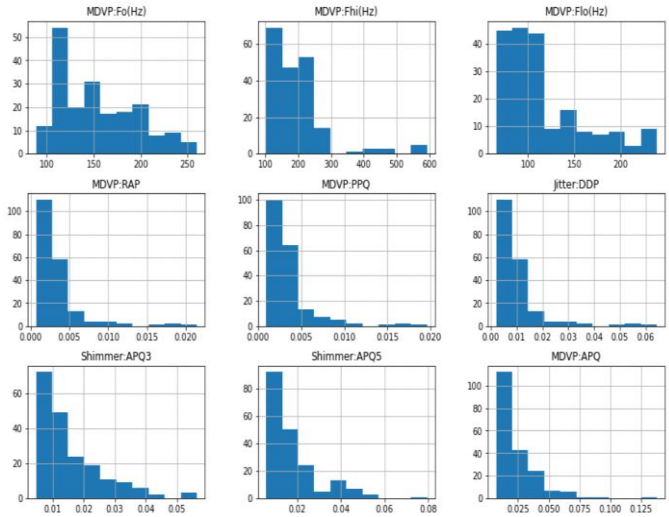

Fig. 6. Distribution Plot for selected Attributes

In Fig. 2, Distribution Plot of MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:RAP, MDVP:PPQ, Jitter:DDP, Shimmer:APQ3, Shimmer:APQ5 and MDVP:APQ are illustrated.

Upon applying the Square-root Function, the skewness of majority of the attributes came down within the permitted range (-1.5<skewness<1.5) and few remaining attributes had their skewness slightly exceeding the range. The Distribution Plot of selected attributes after reducing the skewness is illustrated in Fig. 7.

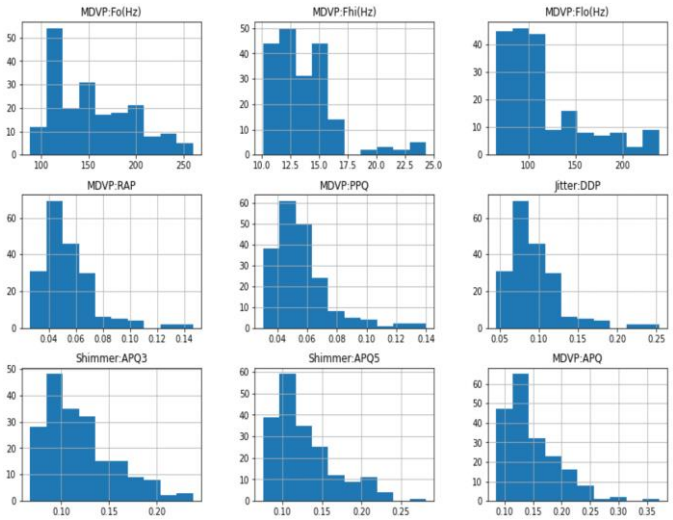

Fig. 7. Distribution Plot for selected Attributes after reducing skewness

In Fig. 3, Distribution Plot of MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:RAP, MDVP:PPQ, Jitter:DDP, Shimmer:APQ3, Shimmer:APQ5 and MDVP:APQ are illustrated after reducing the skewness within the permitted range.

Final check performed on the dataset was for outliers. The presence of outliers can significantly affect the accuracy of the implemented Machine Learning Models. No outliers were

detected in out working dataset. The Pre-processing stage is illustrated in Fig. 8.

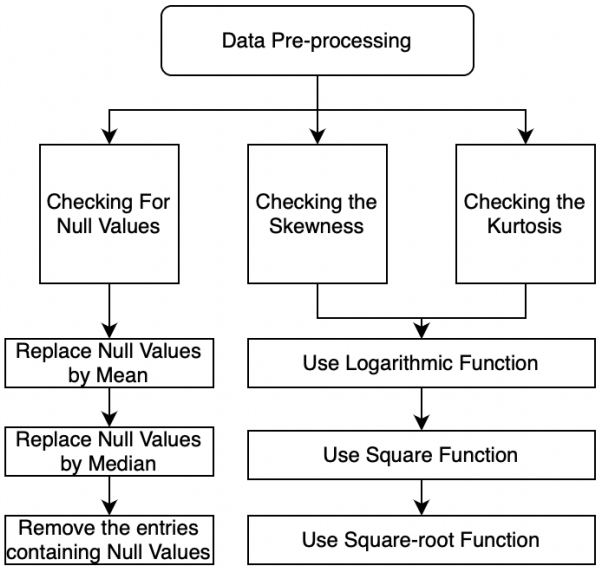

Fig. 8. Data Pre-processing stages followed

C. Feature Selection

One of the most powerful methods to boost the accuracy of the models to a great extent is Feature Selection [16]. Majority of the dataset contains a large number of attributes to work with. Upon close inspection, it is sometimes revealed that not all attributes are useful to the same extent. Including a certain group of attributes, boosts the accuracy of the ML model, while including another group of attributes harms the model. From this observation, it is proved that each attribute has its own weightage. If the ML Models are trained with higher weightage attributes while ignoring the lower ones, then the result obtained will be phenomenal. Two of the most famous Feature Selection Techniques namely, Extra Trees Classifier [17], [18] and Mutual Information [19], [20] are used for the implementation of the ML Models. A brief description is given below.

1) *Extra Trees Classifier (ETC)*

It is a type of ensemble learning technique, which utilizes the concept of decision trees and thus have a resemblance with Random Forest Classifier. Each decision tree is built from the input labelled dataset. At each parent node of the tree, Extra Tree Classifier needs to select the highest priority attribute to continue building. The feature selection is based on a mathematical property called Gini Index. The value obtained from each Gini Index is labelled as Gini Importance of the attribute. After completion of the process, Extra Trees Classifier returns a Gini Importance for each attribute present in the dataset. In Table II the top 11 attributes are chalked out for reference.

TABLE II
GINI IMPORTANCE OF TOP 11 ATTRIBUTES

ATTRIBUTES	GINI IMPORTANCE
spread1	0.112901
MDVP:Fho(Hz)	0.087827
PPE	0.077639
MDVP:Flo(Hz)	0.063037
MDVP:Jitter(Abs)	0.061157
spread2	0.051622
MDVP:Fhi(Hz)	0.049861
D2	0.042425
DFA	0.038613
Shimmer:DDA	0.037636
MDVP:APQ	0.037288

2) *Mutual Information (MI)*

It utilizes the concept of Probability and Information Theory to evaluate the mutual information between the two attributes of the dataset and the process continues iteratively for the rest of them. In a reduced form factor, the Mutual Information of two attributes, say, A and B is given by $MI(A, B) = H(A) - H(A \mid B)$, where MI (A, B) is the mutual information for A and B, H(A) is the entropy of A and H (A | B) is the conditional entropy of A, B. The Mutual Information of the top 11 attribute is portrayed in Table III.

TABLE III
MUTUAL INFORMATION VALUE OF TOP 11 ATTRIBUTES

ATTRIBUTES	MUTUAL IMPORTANCE
PPE	0.253982
spread1	0.220380
MDVP:Fo(Hz)	0.206593
spread2	0.192002
MDVP:Jitter(Abs)	0.191842
MDVP:APQ	0.179583
MDVP:Flo(Hz)	0.174088
MDVP:Fhi(Hz)	0.141811
HNR	0.132462
MDVP:Shimmer(dB)	0.120363
Shimmer:APQ5	0.114313

The steps involved in the Feature Selection Model is illustrated in Fig. 9.

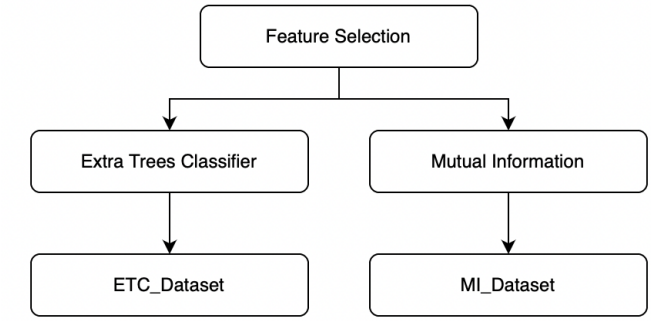

Fig. 9. Feature Selection Process Involved

D. Splitting the Dataset

After performing the Feature Selection, we have obtained two different priority values of the attributes using Extra Trees Classifier and Mutual Information. For Extra Trees Classifier the following attributes have been selected:

- spread1
- MDVP:Fo(Hz)
- PPE
- MDVP:Flo(Hz)
- MDVP:Jitter(Abs)
- spread2
- MDVP:Fhi(Hz)
- D2
- DFA
- Shimmer:DDA
- MDVP:APQ

A Sub-Dataset containing the above-mentioned attributes along with the status column have been created and labelled as ETC_Dataset.

For Mutual Information the following attributes have been selected for use:

- PPE
- spread1
- MDVP:Fo(Hz)
- spread2
- MDVP:APQ
- MDVP:Flo(Hz)
- MDVP:Fhi(Hz)
- MDVP:Jitter(Abs)
- HNR
- NHR
- Shimmer:APQS

A Sub-Dataset containing the above-mentioned attributes along with the status column have been created and labelled as MI_Dataset.

After the creation of ETC_Dataset and MI_Dataset, they are divided into two parts namely, Training Set and Testing Set. The splitting of dataset is in the ratio of 85:15, implying the Training Set to comprise of 1016 cases and Testing Set comprising of 179 cases. For a better understanding, the process is illustrated in Fig. 10.

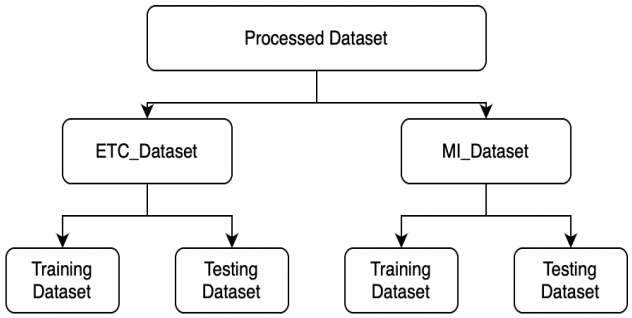

Fig. 10. Splitting of the two dataset

E. Implementation of Machine Learning Models

The following State-Of-The-Art Machine Learning Models are implemented on both the Sub-Datasets:

- K-Nearest Neighbour
- Logistic Regression
- Random Forest Classifier
- XGBoost Classifier

F. ML Model evaluation using performance metrics

The following metrics are used for the performance analysis [21] of all the implemented Machine Learning Models:

- Precision
- Recall
- Accuracy
- F1-Score
- Training Time
- AUC-ROC Curve
- AUC-PR Curve

The steps for deploying a Machine Learning Model includes both the process of Implementation and Model Evaluation using the performance metrics [22]. The workflow is illustrated in Fig. 11 for a clear understanding.

STEPS INVOLVED IN ML MODEL DEPLOYMENT

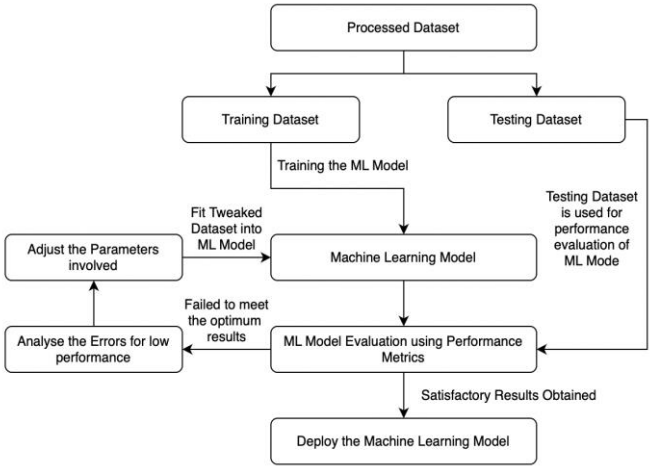

Fig. 11. Steps involved in ML Model Deployment

Precision, Recall, F1-Score and Accuracy are obtained from the Confusion Matrix. It is a 2X2 matrix, containing 4 characteristic values labelled as True Positive (TP), True Negative (TN),

False Positive (FP) and False Negative (FN). Fig. 12 illustrates a typical representation of a Confusion Matrix [23].

TRUE POSITIVE	FALSE POSITIVE
FALSE NEGATIVE	TRUE NEGATIVE

Fig. 12. Confusion Matrix

We can evaluate Precision, Recall, F1-Score and Accuracy in the following manner:

- Precision = (TP) / (TP + FP)
- Recall = (TP) / (TP + FN)
- Accuracy = (TP + TN) / (TP + FP + TN + FN)
- F1-Score = (2 * Recall * Precision) / (Precision + Recall)

VI. Experimental Results And Discussion

A. *Experimental Results obtained for ETC_Dataset*

The True Positive, True Negative, False Positive and False Negative values are portrayed in Table IV.

TABLE IV
VALUES FROM CONFUSION MATRIX FOR ML MODELS

Ml Model	True Positive	False Positive	False Negative	True Negative
KNN	54	12	6	107
LR	48	18	6	107
RFC	54	12	0	113
XGBC	60	6	0	113

From Table IV it is evident, XGBoost Classifier possess the highest value for correctly classifying the patients for Parkinson's Disease and lowest value for misclassification. Utilizing the values obtained from Table IV, Precision, Recall, Accuracy and F1-Score is calculated. The results are tabulated in Table V for a better visualization.

TABLE V
EXPERIMENTAL RESULTS OBTAINED FROM ETC_DATASET

Ml Model	Accuracy	Recall	Precision	F1-Score
KNN	0.9	0.9	0.8182	0.8571
LR	0.8677	0.8889	0.7273	0.7999
RFC	0.9333	1.0	0.8182	0.9
XGBC	0.967	1.0	0.9091	0.9524

From the results of Table V, it is evident that XGBoost Classifier have the highest accuracy of 0.967, followed by Random Forest Classifier with 0.933 and K-Nearest Neighbour with 0.9.

AUC-ROC Curves of implemented ML Models

AUC-ROC Curve stands for Area Under Curve – Receiver Operating Characteristic [24]. It is a very important metric for the performance evaluation of Binary Classification Problems. The value obtained implies how well the ML Model is able to distinguish between the classes. A value closer to 1, means the ML Model has a very high accuracy in differentiation of classes.

The AUC-ROC Curves of the 4 ML Models namely KNN, LR, RFC, XGBC are shown in Fig. 13, 14, 15 and 16 respectively.

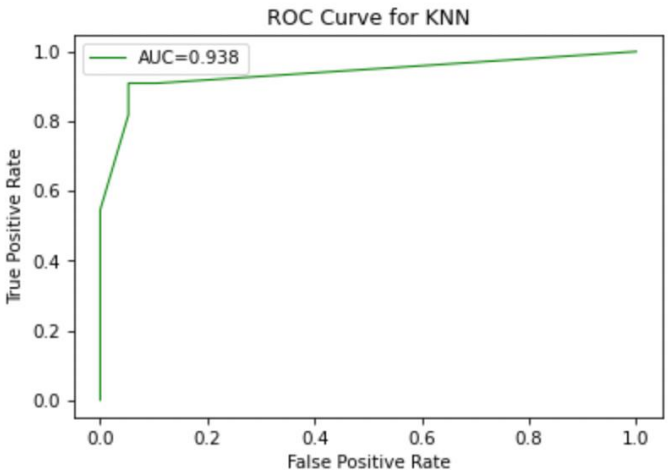

Fig. 13. AUC-ROC Curve for K-Nearest Neighbour

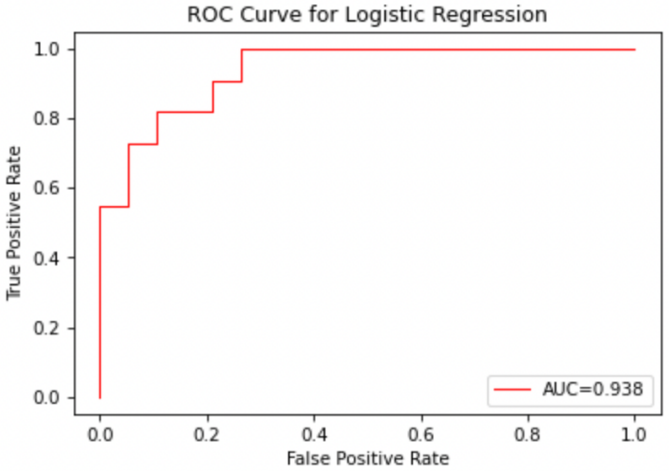

Fig. 14. AUC-ROC Curve for Logistic Regression

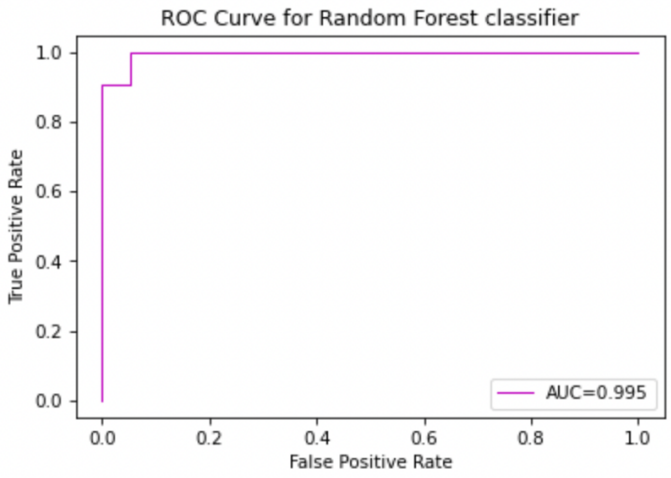

Fig. 15. AUC-ROC Curve for Random Forest Classifier

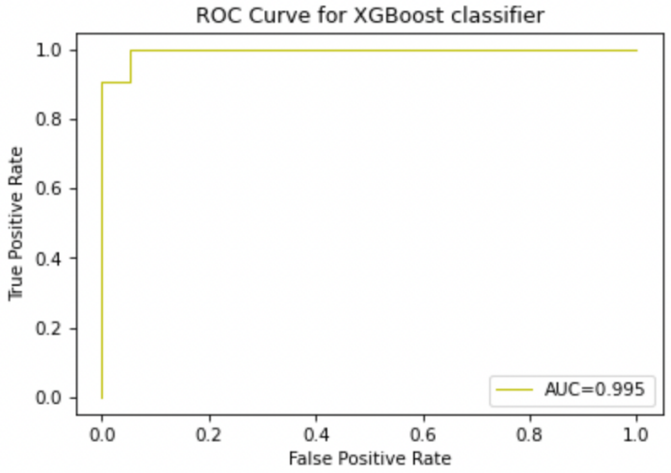

Fig. 16. AUC-ROC Curve for XGBoost Classifier

All the values of AUC-ROC curve are tabulated in Table VI for a compact view.

TABLE VI
PERFORMANCE COMPARISON OF THE ML MODELS BASED ON AUC-ROC VALUE

ML Model	AUC-ROC Value
K-Nearest Neighbour	0.938
Logistic Regression	0.938
Random Forest Classifier	0.995
XGBoost Classifier	0.995

From Table VI, it is clearly showcased that XGBoost Classifier have the highest value of 0.995, followed by Random Forest Classifier with 0.995 and Logistic Regression with 0.938.

AUC-PR Curves of implemented ML Models

AUC-PR stands for Area Under Curve – Precision Recall [25]. The graph is obtained by plotting Precision against Recall for threshold values. An AUC-PR value close to 1 implies, a very good performance from the implemented ML Models

The AUC-PR Curves of the 4 ML Models namely KNN, LR, RFC, XGBC are shown in Fig. 17, 18, 19 and 20 respectively.

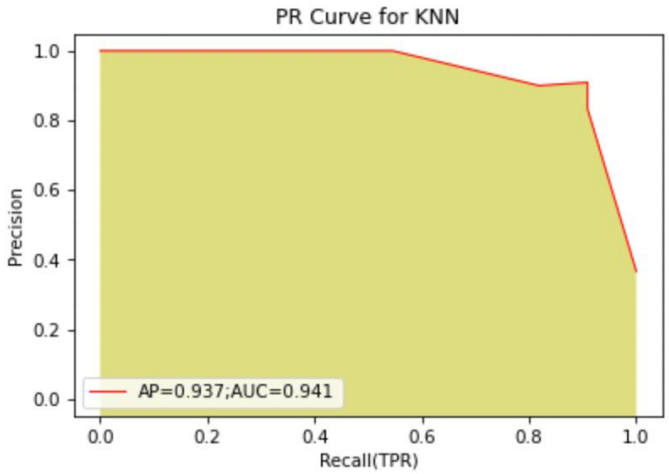

Fig. 17. AUC-ROC Curve for K-Nearest Neighbour

Fig. 18. AUC-ROC Curve for Logistic Regression

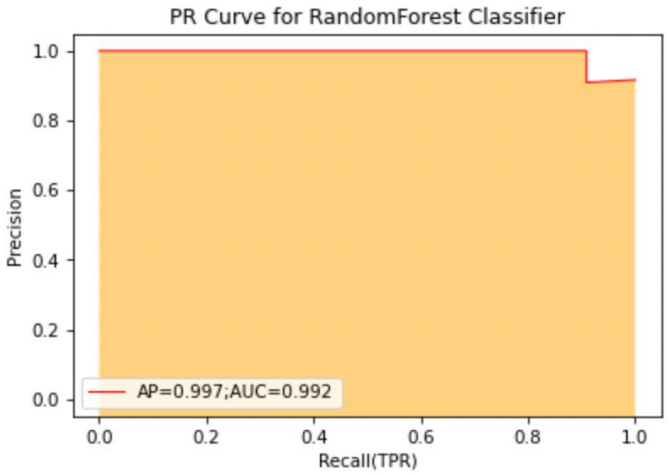

Fig. 19. AUC-ROC Curve for Random Forest Classifier

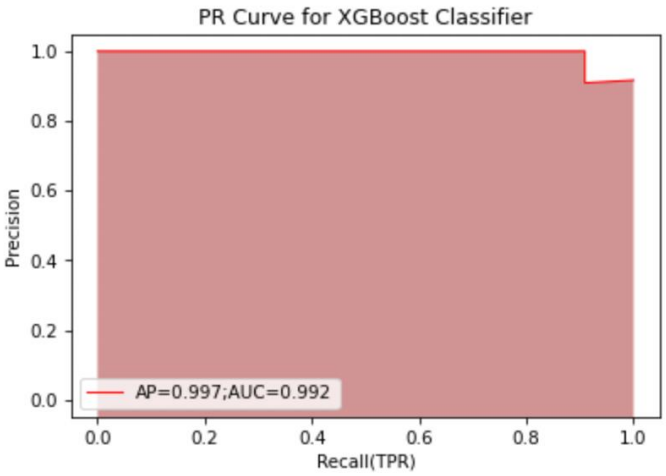

Fig. 20. AUC-ROC Curve for XGBoost Classifier

All the values of AUC-PR and AP are tabulated in Table VII for a compact view.

TABLE VII
PERFORMANCE COMPARISON OF THE ML MODELS BASED ON AUC-PR AND AP VALUE

ML Model	AUC-PR Value	AP Value
K-Nearest Neighbour	0.941	0.937
Logistic Regression	0.903	0.967
Random Forest Classifier	0.992	0.997
XGBoost Classifier	0.992	0.997

From Table VII, it is clearly showcased that XGBoost Classifier have the highest value of 0.992, followed by Random Forest Classifier with 0.992 and K-Nearest Neighbour with 0.941.

B. Experimental Results obtained for MI_Dataset

The True Positive, True Negative, False Positive and False Negative values are portrayed in Table VIII.

TABLE VIII
VALUES FROM CONFUSION MATRIX FOR ML MODELS

Ml Model	True Positive	False Positive	False Negative	True Negative
KNN	60	6	12	101
LR	42	24	6	107
RFC	42	24	0	113
XGBC	42	24	0	113

From Table VIII it is evident, K-Nearest Neighbour possess the highest value for correctly classifying the patients for Parkinson's Disease and lowest value for misclassification. Utilizing the values obtained from Table VIII, Precision, Recall, Accuracy and F1-Score is calculated. The results are tabulated in Table IX for a better visualization.

TABLE IX
EXPERIMENTAL RESULTS OBTAINED FROM MI_DATASET

Ml Model	Accuracy	Recall	Precision	F1-Score
KNN	0.9	0.833	0.9091	0.8696
LR	0.8333	0.875	0.6364	0.7368
RFC	0.867	1.0	0.6364	0.7778
XGBC	0.867	1.0	0.6364	0.7778

From the results of Table IX, it is evident that K-Nearest Neighbour have the highest accuracy of 0.9, followed by Random Forest Classifier and XGBoost with 0.867 and Logistic Regression with 0.8333.

AUC-ROC Curves of implemented ML Model

The AUC-ROC Curves of the 4 ML Models namely KNN, LR, RFC, XGBC are shown in Fig. 21, 22, 23 and 24 respectively.

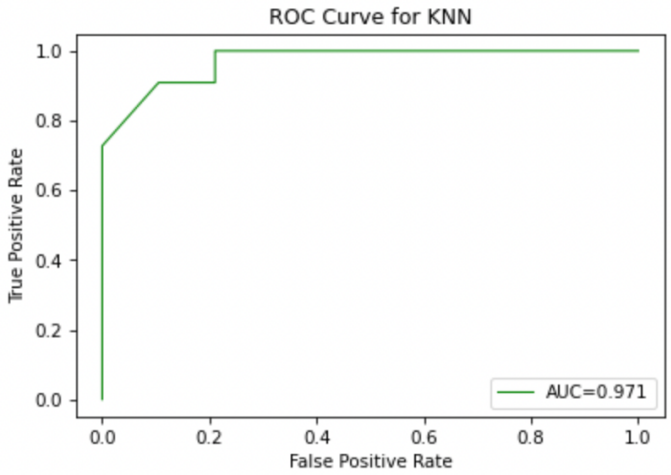

Fig. 21. AUC-ROC Curve for K-Nearest Neighbour

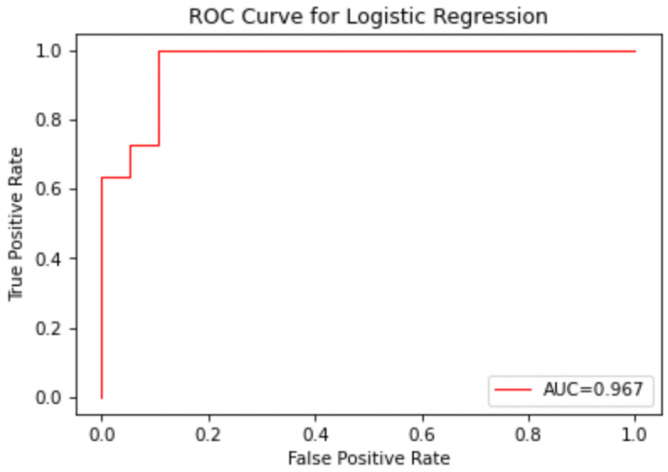

Fig. 22. AUC-ROC Curve for Logistic Regression

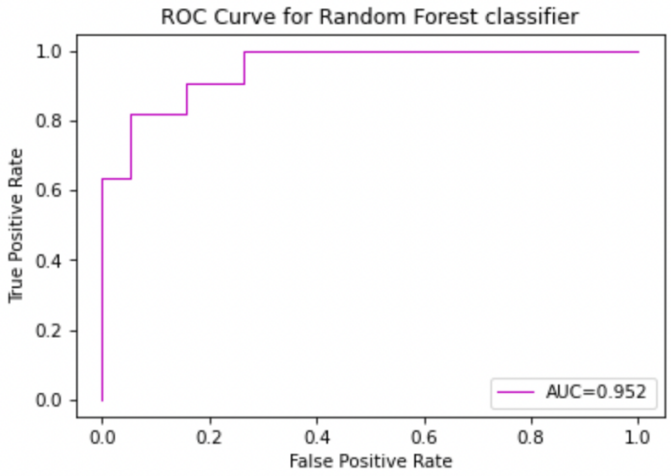

Fig. 23. AUC-ROC Curve for Random Forest Classifier

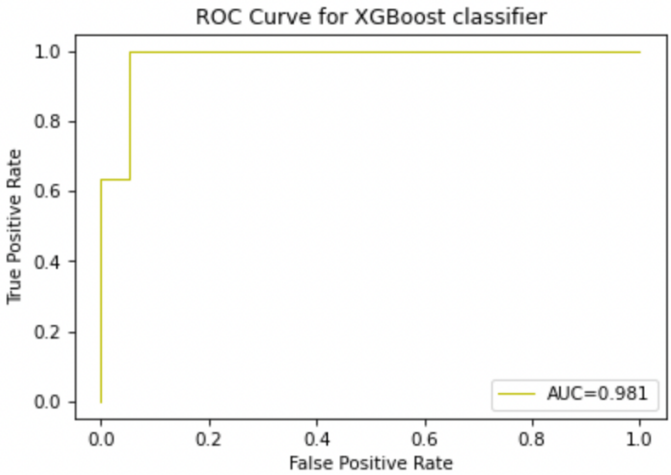

Fig. 24. AUC-ROC Curve for XGBosot Classifier

All the values of AUC-ROC curve are tabulated in Table X for a compact view.

TABLE X
PERFORMANCE COMPARISON OF THE ML MODELS BASED ON AUC-ROC VALUE

ML Model	AUC-ROC Value
K-Nearest Neighbour	0.971
Logistic Regression	0.967
Random Forest Classifier	0.952
XGBoost Classifier	0.981

From Table X, it is clearly showcased that XGBoost Classifier have the highest value of 0.981, followed by K-Nearest Neighbour with 0.971 and Logistic Regression with 0.967.

AUC-PR Curves of implemented ML Models

The AUC-PR Curves of the 4 ML Models namely KNN, LR, RFC, XGBC are shown in Fig. 25, 26, 27 and 28 respectively.

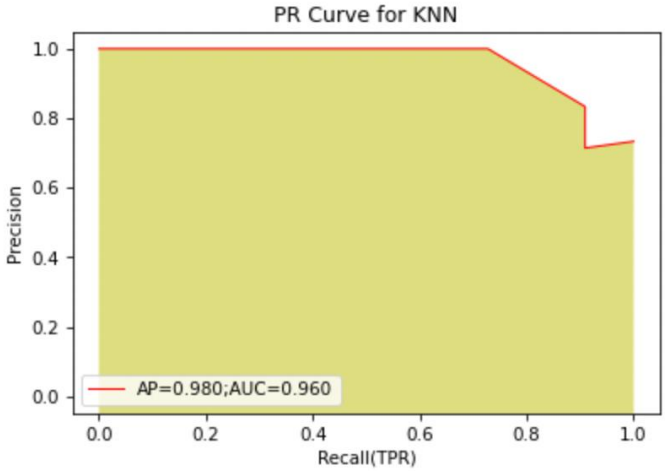

Fig. 25. AUC-PR Curve for K-Nearest Neighbour

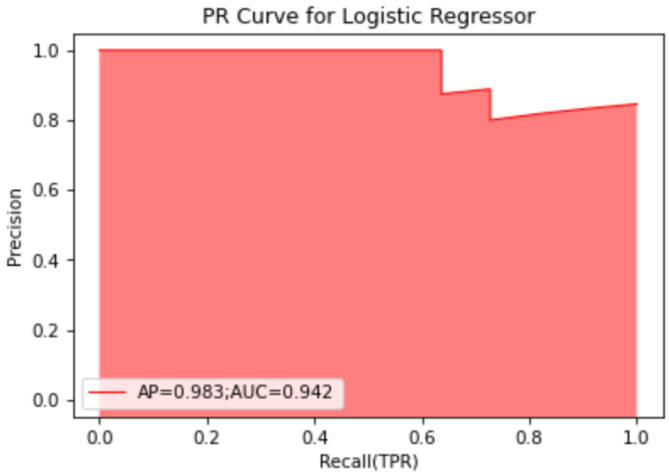

Fig. 26. AUC-PR Curve for Logistic Regression

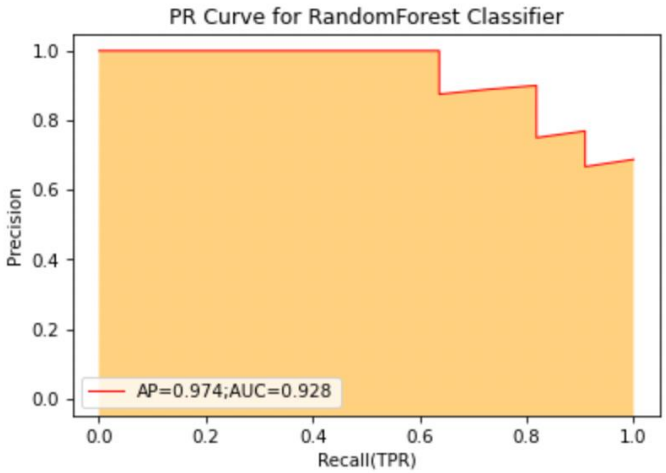

Fig. 27. AUC-PR Curve for Random Forest Classifier

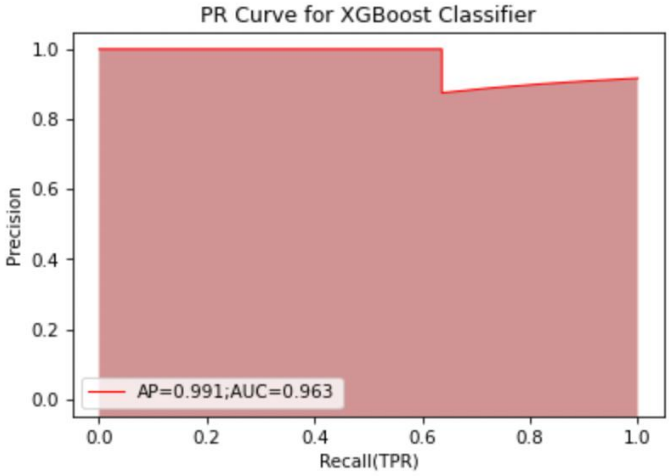

Fig. 28. AUC-PR Curve for XGBoost Classifier

All the values of AUC-PR and AP are tabulated in Table XI for a compact view.

TABLE XI
PERFORMANCE COMPARISON OF THE ML MODELS BASED ON AUC-PR AND AP VALUE

ML Model	AUC-PR Value	AP Value
K-Nearest Neighbour	0.960	0.980
Logistic Regression	0.942	0.983
Random Forest Classifier	0.928	0.974
XGBoost Classifier	0.963	0.991

From Table XI, it is clearly showcased that XGBoost Classifier have the highest AUC-PR value of 0.963 and AP Value of 0.991, followed by K-Nearest Neighbour with AUC-PR Value of 0.960 and AP Value of 0.980.

VII. CONCLUSION AND SCOPE OF FUTURE WORK

The research article focuses on the use of Artificial Intelligence (AI) and Machine Learning (ML) to automate the detection of Parkinson's Disease of a patient. Machine Learning and Artificial Intelligence have helped the Medical Domain to a great extent in reducing diagnosing time and achieving high accuracy rate for correct diagnosis of an ailment.

We have implemented four (04) State-Of-The-Art Machine Learning models namely, K-Nearest Neighbour (KNN), Logistic Regression (LR), Random Forest Classifier (RFC) and XGBoost Classifier (XGBC) to diagnose whether a patient is suffering from Parkinson's Disease or not.

With the use of Feature Selection Technique, the implemented ML Models came up with impressive results. In this article, we have used the two most famous Feature Selection Techniques namely, Extra Trees Classifier (ETC) and Mutual Information (MI).

The results obtained from the implemented ML Models after Data Processing, are evaluated against Precision, Recall,

Accuracy, F1-Score, AUC-ROC Curve and AUC-PR Curve (the performance evaluation metrics).

After the performance evaluation, it is concluded that under the use of Extra Trees Classifier (Feature Selection Technique), XGBoost Classifier came up with the highest accuracy of 0.967, precision of 0.9091, recall of 1.00, F1-Score of 0.9524, AUC-ROC of 0.995, AUC-PR of 0.992 and AP of 0.997.

With the advent of more precise attributes in the future, the accuracy of the ML Models can be boosted to a large extent using the Feature Selection Techniques and effective Data Pre-processing.

REFERENCES

[1] P. Raundale, C. Thosar and S. Rane, “Prediction of Parkinson’s disease and severity of the disease using Machine Learning and Deep Learning algorithm,” *2021 2nd International Conference for Emerging Technology (INCET)*, 2021, pp. 1-5, doi: 10.1109/INCET51464.2021.9456292.

[2] P. Mounika and S. G. Rao, "Machine Learning and Deep Learning Models for Diagnosis of Parkinson’s Disease: A Performance Analysis," *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2021, pp. 381-388, doi: 10.1109/I-SMAC52330.2021.9640632.

[3] A. Ranjan and A. Swetapadma, "An Intelligent Computing Based Approach for Parkinson Disease Detection," *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, 2018, pp. 1-3, doi: 10.1109/ICAECC.2018.8479490.

[4] O. Y. Chén *et al*., "Building a Machine-Learning Framework to Remotely Assess Parkinson's Disease Using Smartphones," in *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 12, pp. 3491-3500, Dec. 2020, doi: 10.1109/TBME.2020.2988942.

[5] L. Zhang, C. Liu and X. Zhang, “Classification of Parkinson’s disease and essential tremor based on structural MRI,” *2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)*, 2016, pp. 410-412, doi: 10.1109/SKIMA.2016.7916256.

[6] N. Jahan, A. Nesa and M. A. Layek, "Parkinson's Disease Detection Using CNN Architectures withTransfer Learning," *2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2021, pp. 1-5, doi: 10.1109/ICSES52305.2021.9633872.

[7] A. Bhan, S. Kapoor and M. Gulati, "Diagnosing Parkinson’s disease in Early Stages using Image Enhancement, ROI Extraction and Deep Learning Algorithms," *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*, 2021, pp. 521-525, doi: 10.1109/ICIEM51511.2021.9445381

[8] A. Ouhmida, O. Terrada, A. Raihani, B. Cherradi and S. Hamida, "Voice-Based Deep Learning Medical Diagnosis System for Parkinson's Disease Prediction," *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1-5, doi: 10.1109/ICOTEN52080.2021.9493456.

[9] K. A. Hasan and M. A. M. Hasan, "Classification of Parkinson's Disease by Analyzing Multiple Vocal Features Sets," *2020 IEEE Region 10 Symposium (TENSYMP)*, 2020, pp. 758-761, doi: 10.1109/TENSYMP50017.2020.9230842.

[10] P. Mounika and S. G. Rao, "Machine Learning and Deep Learning Models for Diagnosis of Parkinson’s Disease: A Performance Analysis," *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2021, pp. 381-388, doi: 10.1109/I-SMAC52330.2021.9640632.

[11] S. Samsani, "An RST based efficient preprocessing technique for handling inconsistent data," *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2016, pp. 1-8, doi: 10.1109/ICCIC.2016.7919591.

[12] A. A. E. Osman, R. A. El-Khoribi, M. E. Shoman and M. A. Wahby Shalaby, "Preprocessing Trajectory Learning Techniques For Robots: A comparative study," *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, 2022, pp. 1412-1415, doi: 10.1109/DASA54658.2022.9765285.

[13] O. Rubenis and A. Matvejevs, "Approximation of Distribution of Log-returns with Normal Inverse Gaussian Process," 2018 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), 2018, pp. 1-3, doi: 10.1109/ITMS.2018.8552949.

[14] Jeong-Hyun Kim, Jong-Hyun Park and Dong-Joong Kang, "Method to improve the performance of the AdaBoost algorithm using Gaussian probability distribution," *2008 International Conference on Control, Automation and Systems*, 2008, pp. 1749-1752, doi: 10.1109/ICCAS.2008.4694511.

[15] A. Ahmad, I. Azmira, S. Ahmad and Nur Ilyana Anwar Apandi, "Statistical distributions of load profiling data," *2012 IEEE International Power Engineering and Optimization Conference Melaka, Malaysia*, 2012, pp. 199-203, doi: 10.1109/PEOCO.2012.6230860.

[16] M. Peker, A. Arslan, B. Şen, F. V. Çelebi and A. But, "A novel hybrid method for determining the depth of anesthesia level: Combining ReliefF feature selection and random forest algorithm (ReliefF+RF)," *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, 2015, pp. 1-8, doi: 10.1109/INISTA.2015.7276737.

[17] A. Safiya Parvin and B. Saleena, "An Ensemble Classifier Model to Predict Credit Scoring - Comparative Analysis," *2020 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, 2020, pp. 27-30, doi: 10.1109/iSES50453.2020.00017.

[18] R. Upadhyay, P. S. Tanwar and S. Degadwala, "Fracture Type Identification Using Extra Tree Classifier," *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2021, pp. 1-6, doi: 10.1109/I-SMAC52330.2021.9640957.

[19] X. Zhao, S. Yang, S. Shan and X. Chen, "Mutual Information Maximization for Effective Lip Reading," *2020 15th IEEE International Conference on*

Automatic Face and Gesture Recognition (FG 2020), 2020, pp. 420-427, doi: 10.1109/FG47880.2020.00133.

[20] R. Kamimura, "Self-Organized Mutual Information Maximization Learning for Improved Generalization Performance," *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015, pp. 1613-1618, doi: 10.1109/SMC.2015.285.

[21] Y. Wang, "Teaching performance evaluation model based on machine learning," *2022 14th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 2022, pp. 843-847, doi: 10.1109/ICMTMA54903.2022.00171.

[22] B. K. D. Prasad, B. Choudhary and B. Ankayarkanni., "Performance Evaluation Model using Unsupervised K-Means Clustering," *2020 International Conference on Communication and Signal Processing (ICCSP)*, 2020, pp. 1456-1458, doi: 10.1109/ICCSP48568.2020.9182368.

[23] F. J. Ariza-Lopez, J. Rodriguez-Avi and M. V. Alba-Fernandez, "Complete Control of an Observed Confusion Matrix," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 1222-1225, doi: 10.1109/IGARSS.2018.8517540.

[24] S. A. Khan and Z. Ali Rana, "Evaluating Performance of Software Defect Prediction Models Using Area Under Precision-Recall Curve (AUC-PR)," *2019 2nd International Conference on Advancements in Computational Sciences (ICACS)*, 2019, pp. 1-6, doi: 10.23919/ICACS.2019.8689135.

[25] K. H. Brodersen, C. S. Ong, K. E. Stephan and J. M. Buhmann, "The Binormal Assumption on Precision-Recall Curves," *2010 20th International Conference on Pattern Recognition*, 2010, pp. 4263-4266, doi: 10.1109/ICPR.2010.1036.