International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

ML-Powered Handwriting Analysis for Early Detection of Alzheimer's Disease

Syed Moinuddin Majid
PG Scholar
Department of CSE
Deccan College of Engineering and Technology
Affiliated to Osmania University
Hyderabad, Telangana
syedmoinuddinmajid3@gmail.com

Dr. Abdul Khadeer (Ass.t Professor)
Department of CSE
Deccan College of Engineering and Technology
Affiliated to Osmania University
Hyderabad, Telangana
abdulkhadeer@deccancollege.ac.in

Abstract: Alzheimer disease (AD) is an irreversible neuronal condition that progressively harms the nerve cells and cognitive functions. Early detection is paramount in enhancing good life since the present methods of treatment are meant to contain its course. One of the fine motor skills impacted by AD may be used as a sign of an early diagnosis, which is handwriting. In spite of its inherent challenges, ML to analyze handwriting has a huge potential in clinical diagnosis. "Analysis of Variance (ANOVA) and Recursive Feature Elimination with Cross-Validation (RFECV)" were employed as feature selection methods in order to identify the most relevant characteristics in the DARWIN dataset. In order to provide credible performance, models were built based on advanced validation methods such as as Monte Carlo Cross-Validation and as Repeated K-Fold. When combined with a number of ML models, a voting classifier reached 88.6% accuracy with RFECV and 100% accuracy with features selected via ANOVA. These results show the effectiveness of feature selection and handwriting analysis procedures in developing reliable diagnostic tools of early Alzheimer disease.

*"**Index Terms -** Alzheimer's disease prediction, ensemble machine learning, handwriting analysis, machine learning for disease prediction".*

I. INTRODUCTION

Dementia is a significant health issue in the world and affects more than 55 million individuals with more than 60 percent of the victims residing in low- and countries with middle-level incomes. Dementia is now the fifth most common cause of death. with approximately 10 million cases being registered annually. It highly contributes to dependency and incapacity particularly to the elderly. In 2019, the economic cost of dementia was more than 1.3 trillion US dollars. Half of this cost was attributed to informal caring where caregivers worked an average number of five hours a day. It is in this light that dementia is a major public health issue that an excessively large amount of money is paid by women due to the fact that they spend the most time taking care of their elderly, and their disability-adjusted life years and mortality rates due to dementia are higher [1].

Neurodegenerative diseases, in particular, AD cause between 60 and 70 percent of diagnoses of dementia. AD causes gradual cognitive decline that begins with impairment of episodic memory due to dysfunction of ventromedial temporal lobes [2]. The worsening of the illness leads to the amnesia and generalized cognitive impairment of the patients, implying the widespread brain damage. Regrettably, AD has no known cure and the primary purpose of the existing drugs is to slow down but not prevent the development of the disease. It is theorized that prevalence of AD will continue to go up as life expectancy increases around the world, and hence, more urgent attention is needed towards developing improved clinical methods of early diagnosis.

During the recent years, motor control activities, including handwriting, have been investigated as a non-invasive approach to the assessment of neurodegenerative diseases because of the interrelation between cognitive and motor functions in the planning and execution of the motions. Regular motor skills are required in handwriting, and changes in the handwriting styles could indicate cognitive and motor impairment related to disorders like AD [3]. To assess motor impairment, which is a symptom associated with neurodegenerative diseases, handwriting analysis through usual graphic tablets make it possible to obtain kinematic and dynamic information, including stroke velocity, pressure and tremors [4], [5], [6]. To evaluate such handwriting traits and develop automation solutions in determining diseases such as AD and Parkinson disease, scientists have turned to ML solutions [7], [8]. These ML-based algorithms can potentially streamline clinical examinations and provide a cost efficient and efficient method of detecting diseases in their initial stages which can be used to complement the existing diagnostic methods [9], [10].

II. RELATED WORK

Early diagnosis is critical in postponing the spread of dementia particularly the AD which remains to be one of the biggest worldwide health issues. Two of the most promising methods of diagnosis of AD that have been highlighted in the recent publications are ML models and neuroimaging. In particular, ML algorithms have become a popular method to use structural MRI, PET, among other types of neuroimaging modalities and obtain more accurate diagnostic tools.

The key focus in the identification of AD is structural MRI that captures the complex brain structure that is often altered in individuals with AD. In a paper by Abbas et al. [11], a transformed domain CNN based on the use of structural MRI data was introduced as a noteworthy paper in the field of AD detection. This paper demonstrated that ANN models could detect subtle changes in the brain that were attributed to Alzheimer disease. Their strategy is to improve the diagnostic potential of MRI scans, which provides a reliable way of detecting AD at the initial stages. Likewise, a method that determines the heterogeneity of brain tissue in AD patients through texture analysis of structural MRI data was also provided by Silva et al. [13]. This technique is crucial in that it enhances detection of AD at an early stage since it reveals microscopic structural changes which can hardly be viewed using the traditional imaging systems.

Besides structural MRI, MRI/PET imaging has also been examined as an alternative that will help improve AD, moderate cognitive impairment (MCI), and healthy aging categorization. To have a more comprehensive view of the brain activity and structure, Rallabandi and Seetharaman [14] applied the DL-based categorization algorithms, which involve combining MRI and PET scans. The relevance of multimodal data in enhancing the accuracy of AD detection was brought to the fore by the level of excellence of their model to distinguish among various stages of cognitive decline.

PET imaging has also been used in anomaly detecting of AD. Baydargil et al. [15] proposed an entropy-based probability model to assess the PET scan to identify abnormalities that are characteristic of AD. Their approach was able to distinguish between the AD patients and the normal controls through the entropy of PET scans in order to identify abnormalities in brain functions. This methodology is consistent with an overall trend of mapping the activity of the brain, often disrupted in neurodegenerative disorders, with functional imaging methods, such as PET.

Also, ML models and genetic data have been integrated to improve early AD detection. Ahmed et al. [16] focused on the early diagnosis by analyzing the genetic variations associated with AD, i.e., single nucleotide polymorphisms (SNPs). Their SNP analysis based on gradient boosting trees revealed that genetic information can provide valuable information on the early development of AD that can be used as an addition to behavioral and neuroimaging data.

A more advanced study has been done on the integration of diagnostic data of multiple sources which include genetic, neuroimaging, and clinical data. To discover biomarkers of AD, Yin et al. [17] used an orthogonal structured sparse canonical correlation analysis method of combining info from imaging and diagnosis. To get a more comprehensive method of diagnosis, their model aimed at enhancing the discovery of specific patterns in their brain scans that interface with genetic markers of AD. This is in accordance with the growing trend of utilizing integrated data sources in order to have a more comprehensive picture of the course of AD.

Transfer ML is another technique that is effective in the diagnosis of AD. In their study, Alatrany et al. [18] also relied on ML to evaluate AD using genome-wide data by using pre-trained models. They have found that DL techniques can be applied to genomic data, and this can enhance the accuracy of AD predictions as it relies on information in other areas. When labelled data is often scarce, as is the case in the context of large scale genomic datasets, this approach is incredibly useful.

Moreover, the research on the less intrusive and more accessible diagnostic tools, like hand written analysis, has shown a great potential in detecting AD. Handwriting analysis requires the integration of cognitive and motor control; one that is often impaired in AD patients. To detect the signs of cognitive decline early on, ML algorithms have been used recently to test the handwriting qualities of strokes velocity, pressure, and tremors. Handwriting also serves as an effective method in the diagnosis of such conditions early because it involves fine motor control and also points to the deterioration of the motor skills which is common in neurodegenerative diseases.

Besides the traditional imaging methods, handwriting analysis is a cheap and non-invasive method of assessing motor function. ML algorithms may be applied to kinematic and dynamic data of digital writing assignments, such as, to generate quantitative measurements, which are correlated with cognitive decline. The approach has the potential to be used as a screening device of

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

Alzheimer due to its ability to be conducted fast and cheaply using the technology that is easily accessible such as graphic tablets. Having focused on the role of handwriting as a biomarker of cognitive impairment, scholars have begun exploring the ways in which changes in handwriting can be used as an early indicator of AD.

The numerous diagnostic methods, which include handwriting analysis, neuroimaging, and many others, are likely to ultimately lead to an early and more accurate diagnosis of AD. A blend of the methods may provide a solid solution to the challenges of the early AD diagnosis due to the ability of ML to analyze large amounts of data. As an instance, integrating the structural and functional imaging data with ML algorithms to evaluate the changes in the handwriting patterns could help in the creation of a more comprehensive diagnostic paradigm, which does not only help in identifying AD early but also tracking its progression over time.

Further, as illustrated by Abbas et al. [11] and Silva et al. [13], there is a lot of potential to enhance the accuracy of diagnosis of neuroimaging data through the use of CNNs and DL models. Such models prove to be valuable in a clinical and research setting since they are capable of identifying patterns that are otherwise too subtle to the human eye with the benefit of exposure to large datasets. A combination of DL and advanced imaging techniques can entirely change the way physicians detect and manage the Alzheimer disease, and it may result in earlier intervention and a higher success rate of the patients.

III. MATERIALS AND METHODS

The aim of the proposed strategy is to develop an analysis of early Alzheimer-based identification of AD using handwriting as the input and ML as the framework. The system will use the DARWIN to determine the notable characteristics that reflect cognitive and motor deficits associated with AD. In order to make sure that the most relevant features will be selected to train the model, two choices of features selection will be utilised Recursive Feature Elimination with "Cross-Validation (RFECV) [12] with Random Forest (RF) and ANOVA [19] (Analysis of Variance)". The system will use many different techniques including; RF, LR, Linear Discriminant Analysis, Gaussian Naive Bayes, Extra Trees, XGBoost, KNN, SVM, Multi-layer Perceptrons, and DT to build the models. Performance improvement will be done using ensemble techniques such as a Voting Classifier (Bagging with RF and DT) and a Stack-Model (RFR + ET with GaussianNB) to improve performance. In a bid to ensure a robust and reliable evaluation of the model performance, two methods, repeated K-Fold and Monte Carlo Cross Validation will be conducted.

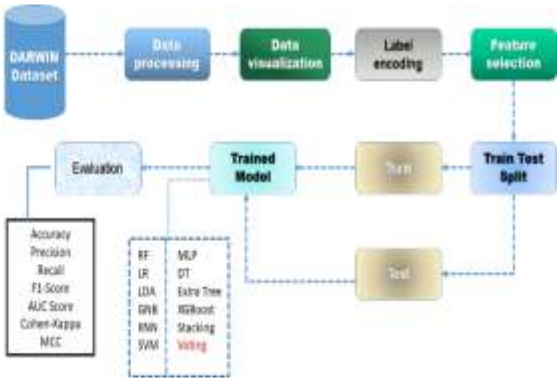


Fig.1 Proposed Architecture

The architecture illustrates a ML pipeline on the DARWIN data. It consists of feature selection, label encoding, data processing and display. Next, the data is split into a training set and a test set. RF, MLP, LR, DT, GNB, XGB, KNN, and SVM are some of the classifiers used to learn the training data. The test set is used to rate the trained models with the AUC-ROC, Cohen Kappa, F1-score, MCC, precision, recall, and accuracy measures.

i) Dataset Collection:

This investigation used the DARWIN dataset that comprises of a range of attributes related to handwriting analysis. It has 25 time points on each feature, among them such attributes as ID, air time 1, disp index 1, gmrt in air 1, gmrt on paper 1, max x extension 1, max y extension 1, mean acc in air 1, and mean acc on paper 1, and so on. class, which identifies the detection of Alzheimer disease, is also represented in the dataset [9]. These features define the kinematic and dynamic features of the handwriting activity.

	ID	air_time1	disp_index1	gmrt_in_air1	gmrt_on_paper1	max_x_extension1
0	id_1	5160	0.000013	120.804174	86.853334	957
1	id_2	51980	0.000016	115.318238	83.448681	1694
2	id_3	2600	0.000010	229.933997	172.761858	2333
3	id_4	2130	0.000010	369.403342	183.193104	1756
4	id_5	2310	0.000007	257.997131	111.275889	987

"Fig.2 Dataset Collection Table – DARWIN"

ii) Pre-Processing:

The pre-processing stage involves preparing the dataset to proceed to a modeling stage. This involves the cleansing of data, representation of significant relationships, coding of nominal labels and feature selection to ensure the high quality input to the prediction model.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

a) Data Processing: Data cleaning is important in order to address data gaps, discrepancies, or errors. This stage is to find and repair any null or duplicate values in the dataset. In order to reduce noise and complexity in computation, the unwanted columns, including IDs or unwanted features, are removed. The features which are not involved in predictive modeling include the ID and other such features. This ensures that the dataset will never prove to be very long and irrelevant to further studies. Clean data reduces overfitting and improves the performance of a model by improving generalization.

b) Data Visualization: Data visualization is of immense help in understanding the feature interactions and in gaining overall insight into a dataset. The correlations between the different aspects can be visualized as this was used to identify the highly linked variables such as multicollinearity was detected using the correlation matrix. The assistance of this phase helps to find redundant characteristics that might affect the model correctness. Sample outcome visualization also helps in understanding the model in response to the various data inputs. Visualizing class distribution and feature relevance can also be used to make the choice of model and feature engineering.

c) Label Encoding: Label encoding is used to transform a category value to a numerical value. The stage is necessary because ML algorithms typically require numerical inputs to train the model. Indicatively, the number code 0 (No) and 1 (Yes) was coded into a categorical variable, i.e. Class indicating the Alzheimer disease groups. Label encoding helps process categorical variables properly and convert them into a format, which is sufficient to be used by machine learning models. It enhances training performance and raises the ability of the model to work well with categorical input.

d) Feature Selection: The feature selection would improve the model performance by keeping only the most significant features and eliminating noise and overfitting. RF is used to determine the best features by successively eliminating features that are not as important based on the model performance using RFECV[19]. ANOVA (Analysis of Variance) is also used to find out how statistically significant each factor is in explaining the goal variable. Both methods optimize the accuracy of the prediction that the model makes by choosing the most useful features.

iii) Training & Testing:

The training and testing step includes preparation of data to be used to create and assess the model. It is the pre-processed dataset that is used to train ML models but gives them a chance to find patterns and connections between features and the goal variable. To reduce the amount of mistake, the models change their internal settings while they are training. Once the training is done, there is another set of data that is used to check how well the models work. This ensures that the models are also effective in application when used on new, untested data that informs us about the predictive ability of the models.

iv) Algorithms:

RF: The RF ensemble learning method is used to build several decision trees that are processed more with an aim of improving accuracy and reducing overfitting. It uses the average prediction of a number of trees to obtain credible results and is particularly applicable in dealing with large data sets having numerous attributes.

LR: A logistic regression linear model is used when there are only two categories. It uses a logistic function to figure out the probability of a binary result, which can help when modeling the relationship between a categorical dependent variable and at least one independent variable.

The simple linear regression line,

^y=a+bx y ^ = a + b x (1)

can be interpreted as follows:

A is the intercept which will give the position in which the regression line will cross the x-axis, b is the change in y, which will occur per unit change in x and y is a projected value.

LDA: Linear discriminant analysis is a way of dimensional reduction which serves as a classifier too. LDA is practical in multi-class scenarios as it maximizes the distance between the classes as it tries to find a linear combination of features that best tells the goal classes apart.

GaussianNB: If you use the Bayes theorem and assume that the traits follow a normal distribution, Gaussian Naive Bayes can be used as a probabilistic classifier. It does a good job of sorting things into groups based on the likelihood of each group given a set of characteristics, so it's very useful for working with big datasets that use continuous features.

ExtraTrees: The extra trees ensemble algorithm randomly splits data and builds many trees and averages their outcomes. It has shorter training times compared to other ensemble methods such as Random Forest and it is accurate because it reduces overfitting. It is particularly effective with data sets having complex feature interactions.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362
IJARST

XGB: The powerful gradient boosting model XGBoost will reduce loss by slowly training individuals with weak learning abilities and correcting errors committed in the past. Due to its speed and precision it is often employed in cases of high performance effectively dealing with big datasets whose data is sparse as well as missing some data.

$y\hat{}i=\sum k=1Kfk(xi)$ (2)

Ywhere K is the number of trees in the group, fk(xi), The Kth tree's prediction of the ith data point is fk(xi), and 3iy, which is 3i y3i, is the total projected value of that same data point.

KNN: The KNN is a non-parametric method which involves classifying a data point based on the majority of the neighbors. It is computationally expensive when there are large databases but it is easy and fast especially where the object of decision is irregular.

Euclidean distance of the two points will be calculated. The Euclidean distance of two points in geometry has already been discussed with us. It can be computed as follows:

“dist((x, y), (a, b)) = $\sqrt{(x - a)^2 + (y - b)^2}$ (3)”

SVM: Encouragement A supervised learning model by the name of SVM finds the hyperplane that does the best job of splitting the different classes. It can do this well both in linear and non-linear classification problems, especially in high-dimensional space by determining decision boundaries through the use of kernel functions.

The linear hyperplane equation may be formulated in the following way:

wTx+b=0$wTx+b=0$ (4)

Where:

- ww is the normal vector to the hyperplane, which shows the way that is perpendicular to the plane.
- bb is the bias term, which shows how far the hyperplane is from the origin along the normal vector ww.

MLP: An ANN composed of a series of nodes is known as a multilayer perceptron. It is particularly useful in non-linear classification problems, where the conventional algorithms would have failed due to interaction between different features, and it is capable of complicated patterns using backpropagation.

The MLP networks consist of a lot of interrelated functions. A network would be comprised of three layers or functions.

“f(x) = f (3)(f (2)(f (1)(x))). (5)”

Each of these levels has units that do the affine change of a linear sum of inputs.

DT: A DT is a tree-like representation that is utilized in problems that involve both regression and classification. It splits the data based on the values of features to be able to offer alternative decision paths. It is effective with problems that have non-linear relationships between the description and the goal variable.

Once you know how much an event costs and how likely it is to happen, you can find its expected value using this formula:

“Expected value (EV) = (First possible outcome x Likelihood of outcome) + (Second possible outcome x Likelihood of outcome) - Cost. (6)”

Stacked Model: Stacked models make use of the benefits of every classifier by combining the numerous ones to become better predictors. The technique would be to mingle Gaussian Naive Bayes with the use of the random forest and extra trees in order to create a more robust model that would yield higher levels of accuracy as they identify more complex patterns in a wide variety of angles.

Voting Classifier: Voting Classifier uses many models to enhance robustness and uses majority voting to combine the predictions of the models. It enhances the stability of the model and reduces variance through bagging of the RF and DT, and therefore, it is quite useful to improve the performance of the model on challenging or noisy data.

IV. RESULTS AND DISCUSSION

Accuracy: A test's accuracy depends on how well it can tell the difference between sick and healthy cases. To find out how accurate a test is, look at the number of true positives and true negatives in each case. that are being assessed should be calculated. This mathematically can be expressed in the following manner:

$$"Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} (7)"$$

Precision: Precision is used to determine the percentage of cases or samples which are correctly classified among the cases or samples considered to

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

be positives. Thus, the precision can be calculated with the help of the following formula:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} \, (8)$$

Recall: Recall is a measure that is employed in ML to identify the potential of a model to locate all the pertinent instances of a specific class. It gives the details about the degree to which a model can represent the occurrence of the instances of a particular class by calculating the fraction of the correct predicted positive instances (divided by the total number of actual positives).

$$Recall = \frac{\text{TP}}{\text{TP } + \text{ FN}} \, (9)$$

F1-Score: The F1 score is a way to check how accurate a ML model is. It blends a model's recall and precision. The accuracy measures the number of times a model accurately predicted the entire data.

$$F1\,Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100(10)$$

AUC-ROC Curve: The AUC-ROC curve shows how well classification problems are solved at different levels. The True Positive Rate and the False Positive Rate are plotted against each other using ROC. AUC shows how well the model can tell the difference between classes: a higher AUC means a better model.

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \, (11)$$

Cohen Kappa: Cohen's Kappa (K) measures how much two raters (or judges, observers, etc.) agree when they each categorize things on their own. It is also helpful, especially when the choices are subjective and the categories are nominal (meaning there is no natural order).

$$Kappa(k) = \frac{P_o - P_e}{1 - P_e} \, (12)$$

MCC: The Matthews coefficient also known as MCC is a way to measure how well machine learning binary models do. It finds the link between the expected and observed binary results by looking at the four parts of a confusion matrix.

$$"MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN}} \, (13)"$$

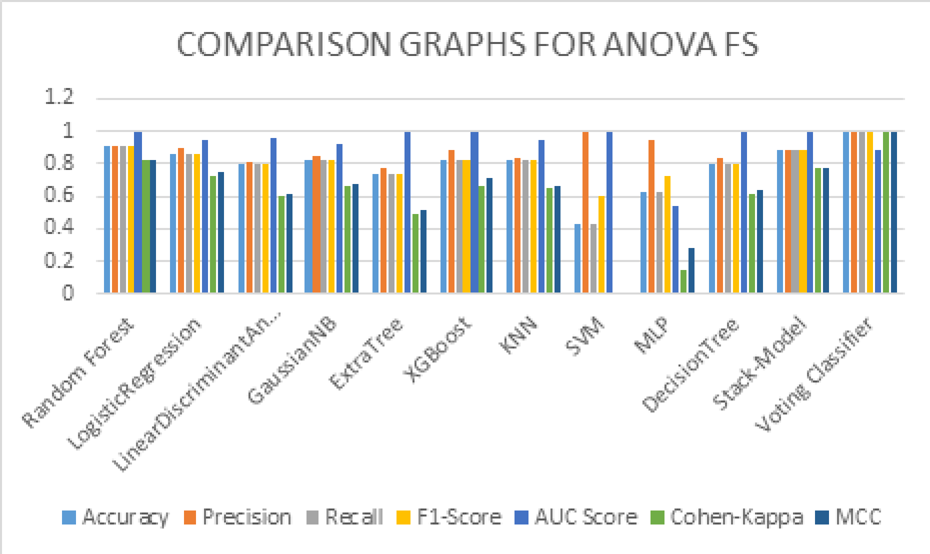
Table 1 and Table 2 measure "the performance of the respective algorithms in terms of accuracy, precision, recall, F1-score, AUC Score, Cohen-Kappa and MCC". The Voting Classifier incessantly beats all the other algorithms in both methods. As well, the tables provide a comparison of the metrics of the different methods.

"Table.1 Performance Evaluation Metrics for Anova FS"

Model	Accuracy	Precision	Recall	F1-Score	AUC Score	Cohen-Kappa	MCC
Random Forest	0.914	0.915	0.914	0.914	1.000	0.826	0.828
LogisticRegression	0.857	0.893	0.857	0.857	0.944	0.720	0.750
LinearDiscriminantAnalysis	0.800	0.810	0.800	0.799	0.953	0.602	0.611
GaussianNB	0.829	0.850	0.829	0.828	0.921	0.661	0.679
ExtraTree	0.743	0.774	0.743	0.743	1.000	0.496	0.517
XGBoost	0.829	0.880	0.829	0.829	1.000	0.667	0.707
KNN	0.829	0.832	0.829	0.828	0.950	0.656	0.660
SVM	0.429	1.000	0.429	0.600	1.000	0.000	0.000
MLP	0.629	0.950	0.629	0.725	0.535	0.150	0.284
DecisionTree	0.800	0.833	0.800	0.800	1.000	0.608	0.633
Stack-Model	0.886	0.890	0.886	0.885	1.000	0.770	0.776
Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**	**0.889**	**1.000**	**1.000**

Graph.1 Comparison Graphs for Anova FS

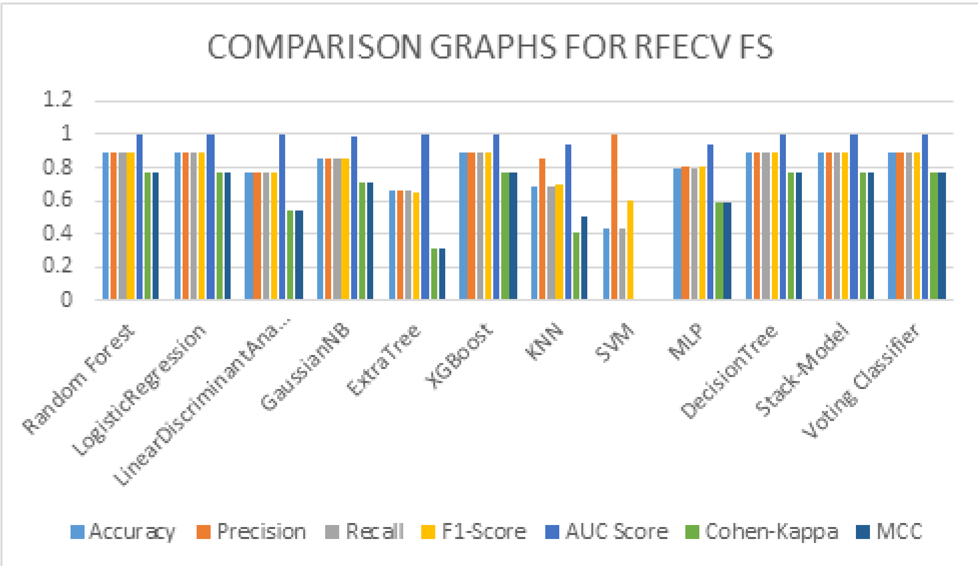
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in



"Table.2 Performance Evaluation Metrics for RFECV FS"

Model	Accuracy	Precision	Recall	F1-Score	AUC Score	Cohen-Kappa	MCC
Random Forest	0.886	0.890	0.886	0.885	1.000	0.770	0.776
LogisticRegression	0.886	0.886	0.886	0.886	1.000	0.767	0.767
LinearDiscriminantAnalysis	0.771	0.774	0.771	0.770	1.000	0.541	0.545
GaussianNB	0.857	0.858	0.857	0.857	0.985	0.711	0.712
ExtraTree	0.657	0.658	0.657	0.655	1.000	0.311	0.314
XGBoost	0.886	0.890	0.886	0.885	1.000	0.770	0.776
KNN	0.686	0.859	0.686	0.703	0.944	0.412	0.510
SVM	0.429	1.000	0.429	0.600	0.000	0.000	0.000
MLP	0.800	0.803	0.800	0.801	0.943	0.588	0.589
DecisionTree	0.886	0.890	0.886	0.885	1.000	0.770	0.776
Stack-Model	0.886	0.890	0.886	0.885	1.000	0.770	0.776
Voting Classifier	**0.886**	**0.890**	**0.886**	**0.885**	**1.000**	**0.770**	**0.776**

"Graph.2 Comparison Graphs for RFECV FS"



In Graphs 1 and 2, F1-Score is light yellow, AUC is blue, Cohen-Kappa is green, MCC is dark blue, accuracy is light blue, precision is orange, recall is grey etc. Voting Classifier is better than the other models in terms of scoring with the two methods as it scores the highest. The graphs above graphically show these results.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

V. CONCLUSION

Overall, the evaluation of various ML methods and feature selection techniques showed a significant difference in performance. The VC (Bagging with RF + DT) has shown a perfect accuracy of 100 percent and this clearly shows the brilliant performance of the feature selection methodology that is based on ANOVA. It shows that the model handles the selected features quite well and makes the underlying patterns of the data to create extremely accurate predictions. Nevertheless, the RFECV method still generated a lower accuracy of 88.6% despite being effective in feature reduction which implies that though RFECV has the ability of reducing the space of features, it may not necessarily lead to optimum performance in this specific problem. The Voting Classifier which was an integration of the benefits of both the RF and DT through bagging was the most effective algorithm in the ones reviewed. These results suggest that when using ANOVA, feature selection should be done carefully in order to maximize model accuracy, and that additional effort in predictive accuracy can be made by using ensemble techniques such as bagging.

To develop the model performance further, the prospective nature of this project will be the investigation of advanced feature selection methods, like RFECV and other feature scoring methods. Also, mixed ensemble methods can be used with DL models to make predictions more accurate. The dataset's spread to a more diverse population and the incorporation of real-time data to continue monitoring the development of Alzheimer disease should contribute to making the dataset more useful and strong in detecting the presence of neurodegenerative disorders at an early stage.

REFERENCES

[1] Dementia. Accessed: Nov. 13, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

[2] M. J. Armstrong, I. Litvan, A. E. Lang, T. H. Bak, K. P. Bhatia, B. Borroni, A. L. Boxer, D. W. Dickson, M. Grossman, M. Hallett, and K. A. Josephs, ''Criteria for the diagnosis of corticobasal degeneration,'' Neurology, vol. 80, no. 5, pp. 496–503, 2013.

[3] R.-E. Precup, T.-A. Teban, A. Albu, A.-B. Borlea, I. A. Zamfirache, and E. M. Petriu, ''Evolving fuzzy models for prosthetic hand myoelectric based control,'' IEEE Trans. Instrum. Meas., vol. 69, no. 7, pp. 4625–4636, Jul. 2020.

[4] G. Vessio, ''Dynamic handwriting analysis for neurodegenerative disease assessment: A literary review,'' Appl. Sci., vol. 9, no. 21, p. 4666, Nov. 2019.

[5] C. De Stefano, F. Fontanella, D. Impedovo, G. Pirlo, and A. S. di Freca, ''Handwriting analysis to support neurodegenerative diseases diagnosis: A review,'' Pattern Recognit. Lett., vol. 121, pp. 37–45, Apr. 2019.

[6] P. Singh and H. Yadav, ''Influence of neurodegenerative diseases on handwriting,'' Forensic Res. Criminol. Int. J., vol. 9, no. 3, pp. 110–114, 2021.

[7] D. Impedovo, G. Pirlo, G. Vessio, and M. T. Angelillo, ''A handwriting based protocol for assessing neurodegenerative dementia,'' Cogn. Com put., vol. 11, pp. 576–586, Aug. 2019.

[8] D. Impedovo, G. Pirlo, and G. Vessio, ''Dynamic handwriting analysis for supporting earlier Parkinson's disease diagnosis,'' Information, vol. 9, no. 10, p. 247, Oct. 2018.

[9] F. Cavallo, A. Moschetti, D. Esposito, C. Maremmani, and E. Rovini, ''Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning,'' Parkinsonism Rel. Disorders, vol. 63, pp. 111–116, Jun. 2019.

[10] M. A. Myszczynska, P. N. Ojamies, A. M. B. Lacoste, D. Neil, A. Saffari, R. Mead, G. M. Hautbergue, J. D. Holbrook, and L. Ferraiuolo, ''Applications of machine learning to diagnosis and treatment of neurodegenerative diseases,'' Nature Rev. Neurol., vol. 16, no. 8, pp. 440–456, Aug. 2020.

[11] S. Qasim Abbas, L. Chi, and Y.-P.-P. Chen, ''Transformed domain convolutional neural network for Alzheimer's disease diagnosis using structural MRI,'' Pattern Recognit., vol. 133, Jan. 2023, Art. no. 109031.

[12] S. Ul Rehman, N. Tarek, C. Magdy, M. Kamel, M. Abdelhalim, A. Melek, L. N. Mahmoud, and I. Sadek, ''AI-based tool for early detection of Alzheimer's disease,'' Heliyon, vol. 10, no. 8, Apr. 2024, Art. no. e29375.

[13] J.Silva, B.C.Bispo, and P.M.Rodrigues, ''Structural MRI texture analysis for detecting Alzheimer's disease,'' J. Med. Biol. Eng., vol. 43, no. 3, pp. 227–238, 2023.

[14] V. P. S. Rallabandi and K. Seetharaman, ''Deep learning-based classifica tion of healthy aging controls, mild cognitive impairment and Alzheimer's disease using fusion of MRI-PET imaging,'' Biomed. Signal Process. Control, vol. 80, Feb. 2023, Art. no. 104312.

[15] H. B. Baydargil, J. Park, and I. F. Ince, ''Anomaly-based Alzheimer's disease detection using entropy-based probability positron emission tomography images,'' ETRI J., pp. 1–13, Mar. 2024.

[16] H. Ahmed, H. Soliman, and M. Elmogy, ''Early detection of Alzheimer's disease using single nucleotide polymorphisms analysis based on gradient boosting tree,'' Comput. Biol. Med., vol. 146, Jul. 2022, Art. no. 105622.

[17] W. Yin, T. Yang, G. Wan, and X. Zhou, ''Identification of image genetic biomarkers of Alzheimer's disease by orthogonal structured sparse canonical correlation analysis based on a diagnostic information fusion,'' Math. Biosciences Eng., vol. 20, no. 9, pp. 16648–16662, 2023.

[18] A. S. Alatrany, W. Khan, A. J. Hussain, J. Mustafina, and D. Al-Jumeily, ''Transfer learning for classification of Alzheimer's disease based on genomewidedata,''IEEE/ACMTrans.Comput.Biol.Bioinf.,vol.20 ,no.5, pp. 2700–2711, May 2023.

[19] B. Jiao, R. Li, H. Zhou, K. Qing, H. Liu, H. Pan, Y. Lei, W. Fu, X. Wang, X. Xiao, X. Liu, Q. Yang, X. Liao, Y. Zhou, L. Fang, Y. Dong, Y. Yang, H. Jiang, S. Huang, and L. Shen, ''Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology,'' Alzheimer's Res. Therapy, vol. 15, no. 1, pp. 1–14, 2023.

[20] S. Qiu et al., ''Multimodal deep learning for Alzheimer's disease dementia assessment,'' Nature Commun., vol. 13, no. 1, p. 3404, 2022.