AI-Driven Graphometric Assessment for Early Cognitive Decline Detection

S Bhanu Teja [1], T Sunil Kumar Reddy [2], G Viswanath[3*], K Yatheendra [4]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
E-mail: 294bhanu@gmail.com ORCID-ID: 0009-0002-2268-1743

[2]Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
E-mail:sunilreddy.vit@gmail.com ORCID-ID: 0000-0003-1257-7985

[3]Associate Professor, Dept. of AI & ML, Sri Venkatesa Perumal College of Engineering & Technology, Puttur
E-mail: viswag111@gmail.com ORCID-ID: 0009-0001-7822-4739

[4]Assistant Professor, Department of AI &ML, Sri Venkatesa Perumal College of Engineering & Technology,Puttur,
E-mail: k.yatheendra84@gmail.com ORCID-ID: 0009-0003-1382-8587

Area/Section: Engineering and Technology
Type of the Paper: Regular Paper
Type of Review: Peer Reviewed as per |C|O|P|E| guidance.
Indexed in: OpenAIRE.
DOI: https://doi.org/10.5281/zenodo.15743135
Google Scholar Citation: IJMTS

How to Cite this Paper:
Teja, S. B., Reddy, T. S. K., Viswanath, G. & Yatheendra, K.(2025). AI-Driven Graphometric Assessment for Early Cognitive Decline Detection. *International Journal of Management, Technology, and Social Sciences (IJMTS), 10*(1), 334-346. DOI: https://doi.org/10.5281/zenodo.15743135

International Journal of Management, Technology, and Social Sciences (IJMTS)
A Refereed International Journal of Srinivas University, India.

CrossRef DOI: https://doi.org/10.47992/IJMTS.2581.6012.0390

Received on: 18/04/2025
Published on: 26/06/2025

AI-Driven Graphometric Assessment for Early Cognitive Decline Detection

S Bhanu Teja [1], T Sunil Kumar Reddy [2], G Viswanath [3*], K Yatheendra [4]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
E-mail: 294bhanu@gmail.com ORCID-ID: 0009-0002-2268-1743

[2]Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
E-mail:sunilreddy.vit@gmail.com ORCID-ID: 0000-0003-1257-7985

[3]Associate Professor, Dept. of AI & ML, Sri Venkatesa Perumal College of Engineering & Technology, Puttur
E-mail: viswag111@gmail.com ORCID-ID: 0009-0001-7822-4739

[4]Assistant Professor, Department of AI &ML, Sri Venkatesa Perumal College of Engineering & Technology,Puttur,
E-mail: k.yatheendra84@gmail.com ORCID-ID: 0009-0003-1382-8587

ABSTRACT

Alzheimer's disease is a progressive neurodegenerative disorder that gradually deteriorates cognitive function and motor skills, significantly affecting quality of life. Early identification plays a critical role in delaying disease advancement and improving patient outcomes. Graphometric analysis, specifically handwriting-based assessment, offers a non-invasive, cost-effective, and promising diagnostic path for early cognitive decline detection. Attention Deficit (AD)-related impairments in fine motor control are often reflected in handwriting patterns, making them valuable indicators for cognitive monitoring. Leveraging artificial intelligence and machine learning, a comprehensive methodology was implemented to extract, select, and evaluate critical handwriting-based features using the DARWIN dataset. To identify the most informative attributes, Recursive Feature Elimination with Cross-Validation (RFECV) and Analysis of Variance (ANOVA) were employed. These techniques allowed for the reduction of feature dimensionality and enhanced model interpretability. Robust classification models were developed using advanced validation strategies, including Repeated K-Fold and Monte Carlo Cross-Validation, ensuring reliable generalization performance. The classification framework incorporated a Voting Classifier ensemble, integrating outputs from various machine learning algorithms to maximize prediction accuracy. Notably, the Voting Classifier achieved a perfect accuracy score of 100% when trained on ANOVA-selected features, while attaining an 88.6% accuracy with RFECV-selected features. These findings underscore the potential of combining intelligent feature selection methods with ensemble learning approaches for reliable cognitive decline assessment through handwriting analysis. The integration of graphometric indicators and machine learning techniques presents a viable pathway for the development of early diagnostic tools, aiding timely intervention in Alzheimer's progression and improving patient care strategies.

Keywords: Alzheimer's disease prediction, ensemble machine learning, handwriting analysis, machine learning for disease prediction.

1. INTRODUCTION:

Dementia constitutes an escalating global health epidemic, impacting about 55 million individuals globally, with over 60% of those afflicted living in low- and middle-income nations. Yearly, over 10 million new cases are documented, with dementia emerging as the 7th largest cause of death, greatly impacting disability and dependency, especially in the senior demographic. In 2019, the economic cost of dementia passed 1.three trillion US dollars, with 50% ascribed to informal caregiving, even as caregivers devoted a mean of 5 hours each day to care provision. This highlights the significant public health challenge posed by dementia, with women disproportionately affected, as they endure greater disability-adjusted life years and mortality rates from the condition, while also comprising 70% of caregiving hours.

Neurodegenerative issues, together with “Alzheimer’s disease (ad)”, account for a substantial percentage of dementia cases, representing 60–70% of diagnosis. “Alzheimer's disease” is marked through a progressive deterioration of cognitive functions, to begin with manifesting as episodic memory deficits associated with dysfunction inside the ventromedial temporal lobe. Because the sickness advances, individuals have good sized cognitive decline and amnesia, signifying pervasive mind impairment. Unluckily, there is no treatment for “Alzheimer's disease”, and present treatments predominantly focus on decelerating its development in preference to reversing or ceasing it. As international lifestyles expectancy rises, the incidence of Alzheimer's disease is expected to growth, underscoring the vital necessity for superior scientific strategies for early diagnosis.

Current studies have established a connection between “cognitive and motor functions” in the “planning and execution” of motions, prompting the research of motor manage tasks, consisting of handwriting, as a non-invasive method for assessing neurodegenerative disease. Handwriting necessitates precise motor control, and alterations in writing patterns may signify cognitive and motor deterioration related to situations which include “Alzheimer's disease” [3]. Handwriting analysis, performed using trendy graphic pills, allows the acquisition of kinematic and dynamic records, inclusive of stroke speed, stress, and tremors, which can be utilized to assess motor impairments linked to neurodegenerative disease [4], [5], [6]. Therefore, researchers have employed machine learning approaches to observe handwriting traits and create automatic systems for figuring out illnesses including “Alzheimer's disease and Parkinson's disease” [7], [8]. those machine learning-based methodologies gift huge capacity for optimizing clinical tests, turning in a price-powerful and efficient technique to early disease identification that can beautify current diagnostic techniques [9], [10].

2. OBJECTIVES:

The goal is to design an AI-driven diagnostic approach using handwriting features to detect early cognitive decline in Alzheimer's disease through feature selection, classification, and ensemble learning techniques.

(1) To identify and extract meaningful handwriting features from the DARWIN dataset using advanced statistical techniques such as “Recursive Feature Elimination with Cross-Validation (RFECV) and Analysis of Variance (ANOVA)”.

(2) To develop robust “machine learning” models utilizing Repeated K-Fold and Monte Carlo Cross-Validation strategies to ensure consistent classification performance for early Alzheimer’s detection based on handwriting traits.

(3) To enhance prediction accuracy by employing an ensemble Voting Classifier that combines multiple machine learning algorithms, demonstrating its efficacy on features selected through RFECV and ANOVA methods.

3. REVIEW OF LITERATURE/ RELATED WORKS:

Dementia, especially "Alzheimer's disease (ad)", constitutes a major “global health” concern, with early identity being essential for mitigating its advancement. latest research have emphasized several diagnostic methodologies for Alzheimer's disease (advert), with "machine learning (ML)" fashions and neuroimaging techniques emerging as particularly promising. The examination of structural MRI, "positron emission tomography (pet)", and various neuroimaging techniques has won prominence, incorporated with machine learning algorithms to develop more particular diagnostic units.

A number one emphasis on “Alzheimer's disease” analysis is structural MRI, which gives complicated photos of brain structure frequently changed in affected individuals. A notable study by Abbas et al. [11] offered a transformed domain "convolutional neural network (CNN)" for the analysis of “Alzheimer's disease” utilizing structural MRI data. This observe illustrated the capability of sophisticated neural community models in identifying modest cerebral alterations linked to Alzheimer’s disease. Their approach emphasizes enhancing the diagnostic efficacy of MRI scans, presenting a dependable tool for the early detection of “Alzheimer's disease”. Likewise, Silva et al. [13] employed texture evaluation of structural MRI statistics, providing a method that assesses the heterogeneity of mind tissue in “Alzheimer's disease” patients. This technology is vital because it detects microscopic structural alterations that are not comfortably apparent with conventional imaging techniques, for this reason improving the early identity of Alzheimer's ailment.

Alongside structural MRI, the mixing of MRI and pet imaging has been investigated for its capability to decorate the differentiation of healthful aging, "mild cognitive impairment (MCI), and Alzheimer's disease (ad)". Rallabandi and Seetharaman [34] utilized “deep learning” to know class models that combine MRI and pet information, providing a extra holistic perspective of cerebral activity and anatomy. Their version exhibited robust efficacy in differentiating among wonderful levels of cognitive decline, underscoring the significance of multimodal data in improving the precision of “Alzheimer's disease detection”.

Pet imaging has been hired in the detection of anomalies related to “Alzheimer's disease”. Bay argil et al. [15] added an entropy-primarily based possibility model for the analysis of pet images to pick out abnormalities indicative of “Alzheimer's disease”. Their method employed the entropy of pet photos to detect anomalies in mind feature, assisting in the differentiation of “Alzheimer's disease” patients from wholesome controls. This approach corresponds with a wider trend of using functional imaging tools, together with puppy, to display mind activity often impaired in neurodegenerative issues.

“Machine learning” methods had been merged with genetic facts to beautify early “Alzheimer's disease” prognosis. Ahmed et al. [16] focused on early detection the use of the analysis of "single nucleotide polymorphisms (SNPs)", genetic abnormalities frequently linked to "Alzheimer's disease (ad)". Their utility of gradient boosting bushes to evaluate SNP records discovered that genetic factors should yield good sized insights into the early stages of Alzheimer's disease, thereby augmenting neuroimaging and behavioral records.

Advanced research has investigated the integration of diagnostic data from diverse resources, such as genetic, neuroimaging, and scientific data. Yin et al. [17] proposed an orthogonal dependent sparse canonical correlation analysis method that integrated imaging and diagnostic records to find out biomarkers for “Alzheimer's disease”. Their version sought to improve the detection of distinct patterns in brain scans associated with genetic signs of Alzheimer's disease, imparting an extra complete diagnostic method. This corresponds with the increasing trend of utilizing included records resources to have a greater thorough expertise of ad's progression.

Transfer learning has emerged as an impressive device in the area of “Alzheimer's disease” prognosis. Alatrany et al. [18] applied switch gaining knowledge of to categorize Alzheimer's disease using genome-wide facts, the usage of pre-trained models for the analysis of extensive datasets. Their studies illustrated the viability of using deep learning methodologies on genetic data, probably improving the predictive accuracy of “Alzheimer's disease” via harnessing insights from other fields. This technique is particularly effective within the realm of substantial genomic datasets, in which labeled data is frequently restricted.

Furthermore, the investigation of non-invasive and more on hand diagnostic strategies, which includes handwriting analysis, has verified significant capability in Alzheimer's disease diagnosis. Handwriting evaluation necessitates the amalgamation of “cognitive and motor control”, frequently compromised in “Alzheimer's disease” patients. Current research have hired “machine learning algorithms” to observe handwriting characteristics, together with stroke pace, stress, and tremors, to discover early signs of cognitive deterioration. Handwriting, as a pastime requiring fine motor control, indicates the deterioration of motor capabilities typically observed in neurodegenerative illnesses, hence serving as an important tool for early detection.

Handwriting analysis gives a cost-efficient and non-invasive approach for evaluating motor function, probably complementing traditional imaging strategies. “Machine learning” models can be trained to assess kinematic and dynamic data from digital writing activities, yielding quantitative measurements that correspond with cognitive decline. This approach suggests potential as a screening instrument for Alzheimer's, as it can be carried out rapidly and value-effectively with usually on hand era, such as picture tablets. Researchers have initiated investigations into how changes in handwriting styles might also act as an early indicator of “Alzheimer's disease”, highlighting handwriting's potential as a biomarker for cognitive deterioration.

The amalgamation of many diagnostic units, including neuroimaging and handwriting analysis, may also finally provide greater particular and timely identity of Alzheimer's disease. The capacity of machine mastering to examine significant datasets can also yield a comprehensive solution to the difficulties related to early “Alzheimer's disease” analysis. Using system mastering algorithms to evaluate changes in handwriting styles, along with structural and functional imaging data, may facilitate the improvement of an extra comprehensive diagnostic framework that now not simplest identifies Alzheimer's disease at an earlier stage but additionally monitors its progression over time.

Moreover, the application of “deep learning” models and convolutional neural networks, as evidenced by Abbas et al. [11] and Silva et al. [13], exhibits significant potential in improving the diagnostic precision of neuroimaging information. Thru the analysis of extensive records, these fashions can discern patterns that can be imperceptible to human statement, rendering them fundamental instruments in medical and research environments. The integration of “deep learning”with sophisticated imaging strategies has the potential to transform the detection and monitoring of "Alzheimer's disease" by enabling early intervention and enhancing patient outcomes.

Table 1: Literature Survey Comparison Table

Sl. No	Area & Focus of the Research	The result of the Research	Reference
1	Structural MRI with CNN for Alzheimer's brain pattern recognition.	CNN identified subtle brain changes for early Alzheimer's diagnosis.	S. Qasim Abbas, L. Chi, and Y.-P.-P. Chen(2023). [11]
2	Texture analysis of structural MRI to detect tissue heterogeneity.	Texture features improved early detection of microscopic brain changes.	J.Silva,B.C.Bispo,and P.M.Rodrigues (2023) [13]
3	Multimodal MRI-PET imaging combined using deep learning classifiers.	Successfully distinguished cognitive stages with enhanced model accuracy.	V. P. S. Rallabandi and K. Seetharaman (2023) [34]
4	SNP-based genetic analysis using gradient boosting for Alzheimer's prediction.	Genetic markers improved early-stage Alzheimer's identification capability.	H. Ahmed, H. Soliman, and M. Elmogy (2022) [16]
5	Transfer learning on genome-wide data using deep neural models.	Pre-trained networks increased Alzheimer's prediction from genomic datasets.	A. S. Alatrany, W. Khan, A. J. Hussain. Et.al., (2023) [18]

4. MATERIALS AND METHODS:

The suggested approach intends to establish a machine learning framework for the early detection of "Alzheimer's disease (ad)" through handwriting analysis. The system will utilize the DARWIN dataset, concentrating on extracting significant features that reflect cognitive and motor impairments linked to Alzheimer's disease. feature selection could be conducted utilizing methodologies: "Recursive feature elimination with cross-Validation (RFECV) employing Random forest (RF) and ANOVA (analysis of Variance)", as a result making sure the incorporation of the maximum pertinent features for version training. The system will construct models employing several techniques, which includes "Random forest, Logistic Regression, Linear Discriminant analysis, Gaussian Naive Bayes, extra trees, XGBoost, k-Nearest neighbors, support Vector Machines, Multi-layer Perceptron, and selection timber". To enhance overall performance, ensemble techniques such a Stack-model "(Random forest blended with extra bushes and GaussianNB)" and a voting Classifier (Bagging with Random forest and decision Tree) will be hired. Repeated ok-Fold and Monte Carlo move-Validation methods will be utilized to offer a rigorous and dependable assessment of model efficacy.

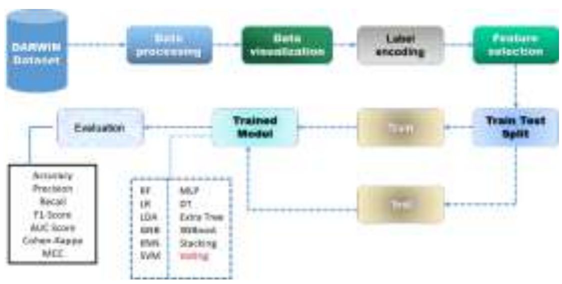


Fig 1: Proposed Architecture

The architecture depicts a "machine learning" pipeline for the DARWIN dataset. It encompasses data processing, visualization, label encoding, and feature selection. The dataset is subsequently divided into training and testing subsets. Multiple classifiers "(RF, MLP, LR, DT, GNB, XGB, KNN, and SVM)" are trained on the schooling dataset. The educated models are assessed at the test set utilizing metrics like as "accuracy, precision, recall, F1-score, Cohen's Kappa, Matthews Correlation Coefficient (MCC), and area under the Receiver operating characteristic Curve (AUC-ROC)".

4.1 Dataset Collection:

This study utilizes the DARWIN dataset, which encompasses several variables pertinent to handwriting analysis. It comprises homes which includes "'identification', 'air_time1', 'disp_index1', 'gmrt_in_air1', 'gmrt_on_paper1', 'max_x_extension1', 'max_y_extension1', 'mean_acc_in_air1', 'mean_acc_on_paper1', among others, totaling 25 time points for each function". The dataset [9] moreover consists of 'class', which signifies the label for "Alzheimer's disease" identification. Those qualities encompass kinematic and dynamic dimensions of handwriting activity.

Table 2: Dataset Collection Table – DARWIN"

	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

4.2 Pre-Processing:

During the pre-processing section, we concentrate on preparing the dataset for modeling. This encompasses data cleansing, showing significant associations, encoding categorical labels, and executing feature selection to guarantee superior input for the predictive model.

4.2.1 Data Processing

Data cleansing is crucial for addressing absent, inconsistent, or erroneous data. This stage entails recognizing and rectifying any null or duplicate values within the dataset. Extraneous columns, including IDs or unnecessary attributes, are eliminated to diminish noise and computational complexity. For instance, attributes such as 'identity' that do not enhance predictive modeling are neglected. This ensures the dataset remains succinct and pertinent for next research. Pristine data enhances model performance by mitigating overfitting and augmenting generalization.

4.2.2 Data Visualization

data visualization is essential for comprehending feature interactions and the general structure of a dataset. The correlation matrix is created to illustrate the correlations among various aspects, aiding within the identification of highly correlated variables and the detection of multicollinearity. This stage assists in detecting superfluous elements that could effect model accuracy. The sample outcome visualization enhances comprehension of the model's conduct with diverse data inputs. Visualizing class distribution and feature significance might inform next choices in model selection and feature engineering.

4.2.3 Label Encoding

Label encoding is employed to convert category values into numerical representations. This stage is essential as machine learning methods normally want numerical inputs for model training. for example, categorical labels such as 'class' denoting Alzheimer's disease categories are converted into numerical values, with 0 representing 'No' and 1 representing 'yes'. Label encoding guarantees the right control and transformation of categorical variables into a layout well suited with machine mastering algorithms. It augments the model's capacity to process categorical data correctly and promotes training performance.

4.2.4 Feature Selection

Feature selection is conducted to preserve only the most pertinent characteristics, enhancing model performance by removing noise and mitigating overfitting. "RFECV (Recursive feature elimination with move-Validation)" in conjunction with "Random forest (RF)" is employed to systematically put off less significant features based on model efficacy, hence identifying the optimal features [19]. Furthermore, ANOVA (analysis of Variance) is utilized to assess the statistical significance of each thing in forecasting the target variable. Each procedures guarantee the selection of the maximum informative characteristics, subsequently enhancing the model's predicted accuracy.

4.3 Training & Testing

Training together with test phase models require a process that prepares data for use in construction as well as evaluation. The pre-developed dataset enables machine learning models to learn both property-target variable patterns and correlations through training processes. The process of training modifies model internal parameters through adjustments which minimize prediction errors. A different dataset is employed for model assessment in further training to determine their performance metrics. Model testing and data normalizing together with predictive skill assessment are guaranteed for these models.

4.4 Algorithms:

RF: Random Forest is access to learning a file that constructs numerous decision -making trees and merges them to increase accuracy and alleviate excess. It is particularly efficient for managing full-size datasets with numerous attributes and yields reliable outcomes by averaging predictions from multiple trees.
LR: Logisticalregress je linear model používaný pro binning klasifikační appliance. It calculates the likelihood of binary results of logistics function, which is effective for modelling the association between a specific variable and one or more independent factors.
LDA: "Linear Discriminant analysis" is a technique for dimensionality reduction that simultaneously functions as a classifier. LDA optimizes class separation by means of identifying a linear combination of features that most effectively distinguishes between target classes, rendering it suitable for multi-class classification challenges.
GaussianNB: "Gaussian Naive Bayes" is a probabilistic classifier that makes use of Bayes' Theorem under the assumption of normally distributed data. It is exceptionally green for huge datasets with continuous variables, offering sincere but a hit type based at the probability of every class given the feature set.
ExtraTrees: ExtraTrees is an ensemble technique that constructs numerous decision trees the usage of random partitions and computes the average in their outputs. It improves accuracy by means of mitigating overfitting and is especially efficacious for datasets with difficult function interactions, supplying expedited training periods relative to other ensemble techniques such as Random forest.
XGB: "XGBoost" is a robust gradient boosting technique that aims to minimize loss through the sequential training of weak learners and the rectification of prior errors. It is often utilized for high-performance applications because to its rapidity and precision, effectively managing enormous datasets with absent values and sparse data.
KNN: "k-Nearest neighbors" is a non-parametric technique that classifies a data item according to the predominant class of its nearest neighbors. It is easy yet efficacious, particularly in scenarios with irregular decision boundaries; yet, it can incur significant processing costs with extensive datasets.
SVM: assistance A Vector machine is a supervised learning model that identifies the hyperplane that optimally distinguishes between several classes. It effectively addresses both linear and non-linear classification challenges, particularly in excessive-dimensional areas, by employing kernel functions to establish decision boundaries.
MLP: A Multilayer Perceptron is a neural network architecture consisting of several layers of nodes. It can learn intricate styles via backpropagation and is especially effective for non-linear classification programs, where conventional algorithms may falter due to feature interactions.
DT: A decision Tree is a hierarchical model utilized for class and regression applications. The dataset is partitioned according to feature values to establish separate decision pathways. It is exceptionally interpretable and efficient for issues involving non-linear correlations between features and the target variable.
Stacked Model: Stacked models integrate several classifiers to enhance predicted accuracy by using the advantages of every. This method integrates Random forest and Extra Trees with Gaussian Naive Bayes to

construct a more robust version, improving accuracy by using identifying intricate patterns from many angles.

Voting Classifier: The voting Classifier employs several models and aggregates their predictions via majority voting to beautify robustness. Using Random forest and decision Tree mitigates version and augments model stability, rendering it very beneficial for enhancing performance on noisy or intricate datasets.

5. RESULTS AND DISCUSSION:

Accuracy: The accuracy of the test concerns its ability to correctly distinguish between patient and healthy cases. In order to assess the accuracy of the test, one must calculate the ratio of real positives and real negatives in all evaluated cases. This can be mathematically expressed as:

“$Accuracy = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}$ (1)”

Precision: The accuracy evaluates the share of precisely classified cases among cases identified as positive. As a result, the formula for calculating accuracy is expressed:

“$Precision = \frac{\text{True Positive}}{\text{True Positive +False Positive}}$ (2)”

Recall: The invocation is a meter in machine learning that evaluates the ability of the model to recognize all relevant cases of a particular class. It is the proportion of precisely predicted positive observations of total real positives and offers insight into the efficiency of the model in identifying the occurrence of a particular class.

“$Recall = \frac{\text{TP}}{\text{TP + FN}}$ (3)”

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$\text{"}F1\ Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100(1)\text{"}$$

AUC-ROC Curve:AUC-ROC serves as an evaluation method that analyzes type performance at multiple threshold levels. ROC graphs the true “positive rate versus the fake positive rate”. Model performance excellence correlates directly with the AUC score because the measure quantifies overall class differentiation ability.

“$AUC = \sum_{i=1}^{n-1}(FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1}+TPR_i}{2}$ (5)"

Cohen Kappa:"Cohen's Kappa (κ)" is a statistical metric employed to assess the degree of concordance between two evaluators who categorize items into distinct classifications. It is particularly advantageous in scenarios when decisions are subjective and the categories are nominal, lacking an herbal order.

“$Kappa(k) = \frac{P_O - P_e}{1 - P_e}$ (6)"

MCC: The Matthews’s coefficient, or "Matthews’s correlation coefficient (MCC)", is a performance indicator applied for binary classifiers in machine learning. It assesses the correlation between predicted and actual binary results by evaluating all 4 components of a confusion matrix.

“$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN}}$ (7)”

Tables (3 & 4)assess the performance metrics—"accuracy, precision, recall, F1-score, AUC score, Cohen's Kappa, and Matthews correlation coefficient—for each algorithm".The voting Classifier often surpasses all different algorithms across each methodologies. The tables provide a comparative exam of the metrics for the alternative methods.

Table 3: 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

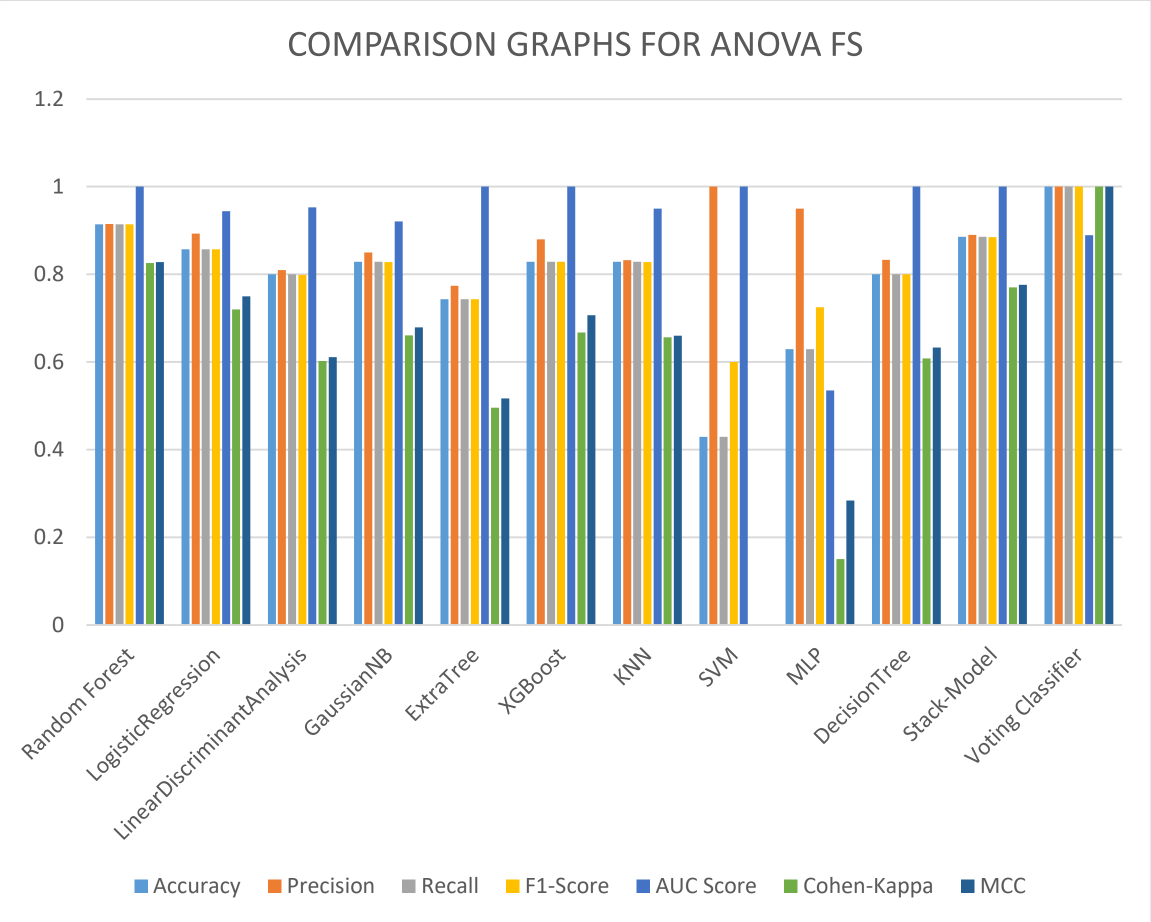
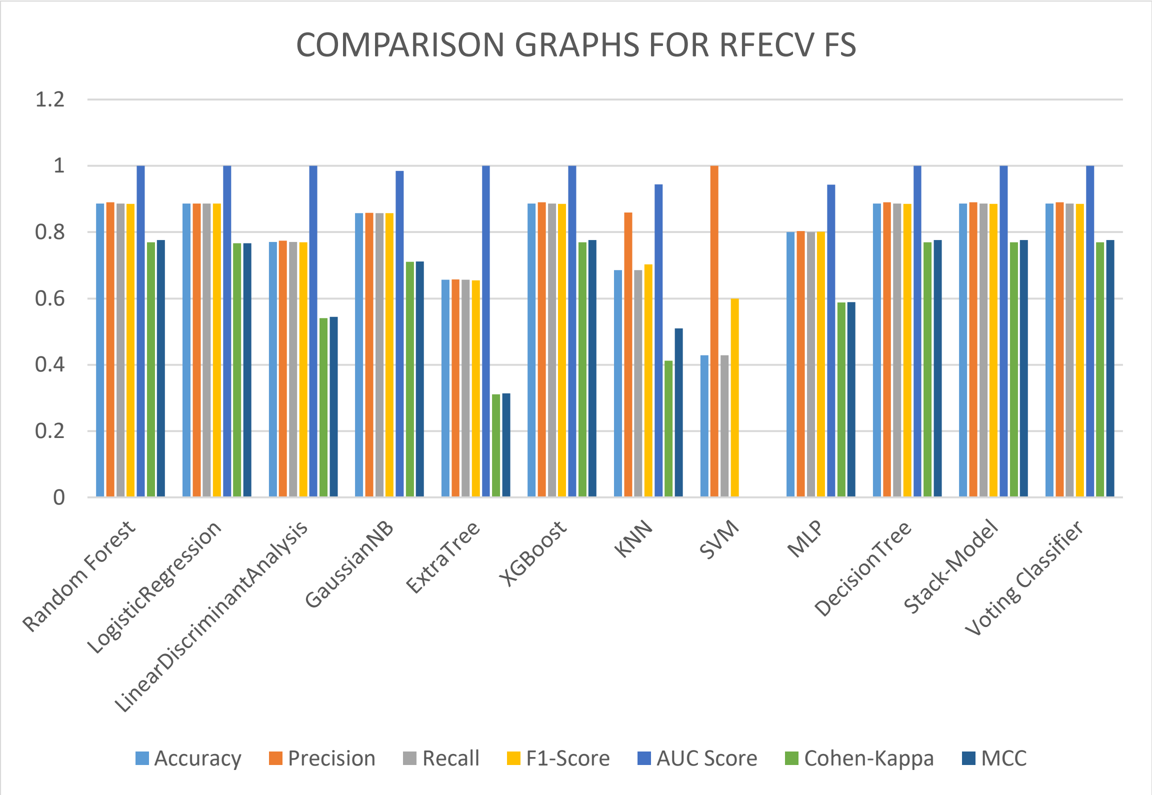


Table 4: Performance Evaluation Metrics for RFECV FS

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



Accuracy is depicted in light blue, “precision in orange, recall in grey, F1-score in light yellow, AUC in blue, Cohen-Kappa in inexperienced, and MCC in dark blue in Graphs 1 and 2”. Relative to the opposite models, the voting Classifier demonstrates enhanced overall performance across both approaches, attaining the highest metrics. The graphs above visually represent these findings.

6. CONCLUSION:

The assessment of diverse feature selection methods and machine learning algorithms revealed substantial performance disparities. The ANOVA-based feature selection method verified exceptional consequences, with the voting Classifier (comprising Bagging with Random forest and decision Tree) attaining flawless accuracy of 100%. This underscores the model's proficiency in managing the selected features adeptly, elucidating the intrinsic patterns inside the data for quite specific predictions. Conversely, the RFECV technique, while talented in characteristic reduction, yielded a faded accuracy of 88.6%, suggesting that although RFECV can condense the characteristic area, it is able to not consistently reap optimal performance for this precise difficulty. The voting Classifier turned into recognized because the simplest model among the algorithms evaluated, utilizing the advantages of both Random forest and decision Tree through bagging. The results indicate that meticulous feature selection, particularly by ANOVA, is essential for strengthening version accuracy, with ensemble methods like as bagging further augmenting predictive performance [20].
The future scope of this work encompasses the investigation of advanced feature selection strategies, including "recursive feature elimination with cross-validation (RFECV)" in conjunction with additional

feature rating techniques, to enhance model performance further. Furthermore, the incorporation of deep learning models and hybrid ensemble strategies may yield advanced forecast precision. Integrating real-time data for ongoing monitoring of Alzheimer's disease development and broadening the dataset to encompass a more diverse demography need to enhance the usefulness and robustness in figuring out early-stage neurodegenerative disorders.

REFERENCES:

[1] Muramatsu, C., Morishita, T., Takahashi, R., Hayashi, T., Nishiyama, W., Ariji, Y., Zhou, X., Hara, T., Katsumata, A., Ariji, E., et al. (2021). Tooth detection and classification on panoramic radiographs for automatic dental chart filing: Improved classification by multi-sized input data. *Oral Radiol*, 37(5), 13–19.

[2] Viswanath, G. (2022). A Smart Recommendation System for Medicine using Intelligent NLP Techniques. *International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 5(2), 1081-1084

[3] Precup, R. E., Teban, T. A., Albu, A., Borlea, A. B., Zamfirache, I. A., & Petriu, E. M. (2020). Evolving fuzzy models for prosthetic hand myoelectric based control. *IEEE Trans. Instrum. Meas.,* 69(7), 4625–4636.

[4] Swapna, G., & Bhaskar, K. (2024). Early-Stage Autism Spectrum Disorder Detection Using Machine Learning. *International Journal of HRM and Organizational Behavior,* 12(3), 269-283, DOI: https://ijhrmob.org/index.php/ijhrmob/article/view/242

[5] Stefano, C. D., Fontanella, F., Impedovo, D., Pirlo, G., & Freca, A. S. D. (2019). Handwriting analysis to support neurodegenerative diseases diagnosis: A review. *Pattern Recognit. Lett.,* 121(2), 37–45.

[6] Viswanath, G. (2024). Multiple Cancer Types Classified Using CTMRI Images Based On Learning Without Forgetting Powered Deep Learning Models. *International Journal of HRM and Organizational Behavior*, 12(3), 243-253.

[7] Impedovo, D., Pirlo, G., Vessio, G., & Angelillo, M. T. (2019). A handwriting based protocol for assessing neurodegenerative dementia. *Cogn. Com put.,* 11(2), 576–586.

[8] Viswanath, G., & Dr. Swapna, G. (2024). Health Prediction Using Machine Learning with Drive HQ Cloud Security. *Frontiersin Health Informatics.* 13(8), 2755-2761.

[9] Cavallo, F., Moschetti,A., Esposito, D., Maremmani, C., & Rovini, E. (2019). Upper limb motor pre-clinical assessment in Parkinson’s disease using machine learning. *Parkinsonism Rel. Disorders,* 63(3), 111–116.

[10] Viswanath, G., & Swapna, G. (2025). Data Mining-Driven Multi-Feature Selection for Chronic Disease Forecasting. *Journal of Neonatal Surgery*, 14(5s), 108-124.

[11] Wang, G., Liu, L., Lauri, F., & El-Hassani, A. H. (2024). Multi-Scale DCNN with Dynamic Weight and Part Cross-Entropy Loss for Skin Lesion Diagnosis. *Big Data Mining and Analytics*, 7(4), 1347-136.

[12] Viswanath, G. (2024). Personalized Breast Cancer Prognosis through Data Mining Innovations. *Cuestiones de Fisioterapia*, 53(2), 538-548.

[13] Silva, J., Bispo, B. C. & Rodrigues, P. M. (2023). StructuralMRItextureanalysis for detecting Alzheimer’s disease. *J. Med. Biol. Eng.,* 43(3), 227–238.

[14] Viswanath, G. (2021). Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage. *Turkish Journal of Computer and Mathematics Education,* 12(9), 545-554.

[15] Zou, L., Luo, X., Zhang, Y., Yang, X., & Wang, X. (2023). HC-DTTSVM: A network intrusion detection method based on decision tree twin support vector machine and hierarchical clustering. *IEEE Access*, 11(1), 21404–21416.

[16] Vachmanus, S., Noraset, T., Piyanonpong, W., Rattananukrom, T., & Tuarob, S. (2023). DeepMetaForge: A Deep Vision-Transformer Metadata-Fusion Network for Automatic Skin Lesion Classification. *IEEE Access.* 11(1), 145467-145484.

[17] Yin, W., Yang, T., Wan, G., & Zhou, X. (2023). 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.,* 20(9), 16648–16662.

[18] Alatrany, A. S., Khan, W., Hussain, A. J., Mustafina, J., & Al-Jumeily, D. (2023). Transfer learning for classification of Alzheimer's disease based on genomewidedata. *IEEE/ACMTrans.Comput.Biol.Bioinf*., 20(5), 2700–2711.

[19] Jiao, B., Li, R., Zhou, H., Qing, K., Liu, H., Pan, H., Lei, Y., Fu, W., Wang, X., Xiao, X., Liu, X., Yang, Q., Liao, X., Zhou, Y., Fang, L., Dong, Y., Yang, Y., Jiang, H., Huang, S., & Shen, L. (2023). Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology. *Alzheimer's Res. Therapy*, 15(1), 1–14.

[20] Qiu, S., et al. (2022). Multimodal deep learning for Alzheimer's disease dementia assessment. *Nature Commun.,* 13(1), 3404, 2022.

[21] Swapna, G. (2023). A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification. *Journal of Computer Science*, 19(3), 1203-1211.

[22] Armstrong, M. J., Litvan, I., Lang, A. E., Bak, T. H., Bhatia, K. P., Borroni, B., Boxer, A. L., Dickson, D. W., Grossman, M., Hallett, M., & Josephs, K. A. (2013). Criteria for the diagnosis of corticobasal degeneration. *Neurology,* 80(5), 496–503.

[23] Viswanath, G. (2024). Machine-Learning-Based Cloud Intrusion Detection. *International Journal of Mechanical Engineering Research and Technology,* 16(5), 38-52.

[24] Onan, A., Bulut, V., & Ezgi, A. (2023). Enhancing Skin Lesion Classification with Ensemble Data Augmentation and Convolutional Neural Networks. *International Congress of Electrical and Computer Engineering,* 2023(1), 131-145.

[25] Swapna, G., & Bhaskar, K. (2024). Malaria Diagnosis Using Double Hidden Layer Extreme Learning Machine Algorithm With Cnn Feature Extraction And Parasite Inflator. *International Journal of Information Technology and Computer Engineering,* 12(4), 536-547.

[26] Singh, P., & Yadav, H. (2021). Influence of neurodegenerative diseases on handwriting. *Forensic Res. Criminol. Int. J.,* 9(3), 110–114.

[27] Viswanath, G., & Swapna, G. (2025). Diabetes Diagnosis Using Machine Learning with Cloud Security. *Cuestiones de Fisioterapia,* 54(2), 417-431.

[28] Ballan, L., Bertini, M., Del Bimbo, A., & Nunziati, W. (2007). Automatic Detection and Recognition of Players in Soccer Videos. *Advances in Visual Information Systems*. *Springer*, 4781(1), 105–116.

[29] Viswanath, G. (2024). Improved Light GBM Model Performance Analysis and Comparison for Coronary Heart Disease Prediction. *International Journal of Information Technology and Computer Engineering,* 12(3), 658-672.

[30] Myszczynska, M. A., Ojamies, P. N., Lacoste, A. M. B., Neil, D., Saffari, A., Mead, R., Hautbergue, G. M., Holbrook, J. D., & Ferraiuolo, L. (2020). Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Rev. Neurol*., 16(8), 440–456.

[31] Viswanath, G., (2024). Enhancing Cloud Security: A Blockchain-Based Verification Framework for Multi-Cloud Virtual Machine Images. *Frontiers in Health Informatics*, 13(3), 9535-9549.

[32] Ellappan, V., & Rajkumar, R. (2021). Classification of cricket videos using finite state machines. *Int. J. Inf. Technol. Manage, 2*0 (1–2), 83- 94.

[33] Viswanath, G. (2021). Hybrid encryption framework for securing big data storage in multi-cloud environment. *Evolutionary intelligence*, 14(2), 691-698.

[34] Mahmood, Z., Ali, T., Khattak, S., Hasan, L., & Khan, S. U. (2014). Automatic player detection and identification for sports entertainment applications. *Pattern Anal. Appl.,* 18(4), 971–982.

[35] Viswanath, G. (2023). A Real-Time Case Scenario Based On URL Phishing Detection Through Login URLS. *Material science and technology,* 22(9), 103-108.

[36] Dhanasekhar Reddy, A., et.al. (2025). Unveiling Hidden Risks in Medication Combinations with Graph-Based Adaptive Learning. *International Journal of Health Sciences and Pharmacy (IJHSP),* 9(1), 96-104. DOI: https://doi.org/10.5281/zenodo.15478353

[37] Naga Raju, D., et.al. (2025). Bio-Inspired Intelligence for Adaptive Risk Assessment in Heart Disease A Synergistic. *International Journal of Health Sciences and Pharmacy (IJHSP),* 9(1), 72-84.

DOI:https://doi.org/10.5281/zenodo.15469250

[38] Sunil Kumar Reddy, T., et.al. (2025). Interpretable AI for Precision Brain Tumor Prognosis: A Transparent Machine Learning Approach. *International Journal of Health Sciences and Pharmacy (IJHSP),* 9(1), 180-195. DOI: https://doi.org/10.5281/zenodo.15523628

****