

EVALUATING MRI PLANE AND MACHINE LEARNING ALGORITHM PERFORMANCE IN ALZHEIMER'S DISEASE CLASSIFICATION USING HARALICK TEXTURE FEATURES

Gayathri L

Department of Computer Science and Applications, Bangalore University, Bangalore, India

Muralidhara B. L

Department of Computer Science and Applications, Bangalore University, Bangalore, India

ABSTRACT:

Classifying Alzheimer's Disease (AD) using MRI scans is essential for timely detection and effective treatment planning. This research aims to enhance AD diagnosis by employing machine learning (ML) models, feature selection methods, and texture-based image analysis. The study compares the effectiveness of various feature selection strategies and Principal Component Analysis (PCA) combined with multiple ML algorithms to determine the most suitable approach for differentiating between Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and AD cases.

The preprocessing workflow includes N4 bias correction, skull stripping, and linear co-registration, followed by the extraction of texture features that capture statistical characteristics of image patterns. Different ML classifiers are then trained and tested on these features to evaluate their ability to accurately categorize patients. Performance is measured using multiple evaluation metrics to assess the discriminative power of the models across AD stages.

The findings highlight that combining ML techniques with feature selection and texture analysis provides a robust framework for early AD detection and personalized treatment strategies, offering meaningful implications for clinical use.

Keywords— Alzheimer's disease, feature selection, machine learning, texture analysis.

1. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative condition characterized by the gradual breakdown of neurons and their connections. This degeneration results in memory decline, impaired cognition, and difficulties in carrying out everyday activities. As the disease advances, the brain undergoes structural changes, including hippocampal atrophy and the enlargement of cerebrospinal fluid-filled spaces [1]. Globally, AD represents a significant public health challenge, impacting millions of individuals and placing heavy economic and social burdens on societies [2]. Currently, over 50 million people are living with dementia worldwide, and this number is projected to rise sharply to approximately 152 million by 2050 [3].

AD progression is generally divided into three stages: Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease. However, achieving reliable diagnosis, especially at the early stage, remains difficult in clinical practice. In neuroimaging, texture-based features derived from Magnetic Resonance Imaging (MRI) play a crucial role in detecting structural variations in brain tissue. These features capture intensity variations and tissue patterns, which help distinguish healthy from diseased brain regions, providing essential input for classification models. To further improve classification accuracy, feature

selection techniques are applied to reduce irrelevant or redundant data. Principal Component Analysis (PCA) is also employed as a dimensionality reduction method, condensing high-dimensional data into a more compact form while retaining essential variance.

Machine Learning (ML) has emerged as a powerful tool for medical image analysis, offering algorithms capable of handling complex datasets and identifying disease-related patterns [6]. Integrating ML with feature selection and dimensionality reduction has the potential to improve classification performance across different AD stages, supporting timely intervention and better patient management.

In this research, we conduct a comparative study of various feature selection methods in combination with PCA and multiple ML algorithms to determine the most effective approach for AD stage classification. For each method, only the top half of the most relevant features are retained. This analysis provides valuable insights into optimizing feature selection and dimensionality reduction strategies for improving AD diagnosis and classification.

2. RELATED WORK

In recent years, computer-aided diagnosis (CAD) systems have become vital in supporting clinicians, particularly in the detection and classification of Alzheimer's disease (AD) stages. Numerous studies have investigated CAD frameworks built on traditional machine learning (ML) methods, integrating texture analysis and feature selection to improve diagnostic accuracy. This section summarizes some of the significant contributions in this area.

The study in [7] introduced a texture-feature extraction framework for volumetric medical imaging. By combining Local Binary Patterns (LBP) with Gray-Level Co-occurrence Matrix (GLCM) descriptors, the method successfully captured key texture attributes such as local anisotropy, contrast, homogeneity, and structural patterns. Compared to conventional 2D/3D features and even deep learning approaches with limited training data, this hybrid approach achieved superior classification outcomes. Similarly, the authors in [8] proposed a feature fusion strategy that integrated GLCM, 3D Scale and Rotation Invariant Feature Transform (3D SIFT), Histogram of Oriented Gradients-Three Orthogonal Planes (HOG-TOP), and Complete LBP of Sign and Magnitude-Three Orthogonal Planes (CLBPSM-TOP). Using the OASIS dataset, they classified participants into CN, MCI, and AD categories via Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and an ensemble model, with the ensemble method demonstrating the best performance.

The work in [9] explored image classification using a Bag-of-Features representation, comparing Speeded-Up Robust Features (SURF) with global color descriptors. Their experiments showed that combining SURF features with a cubic SVM classifier achieved the highest accuracy. In contrast, [6] investigated convolutional neural network (CNN)-based feature extraction for automated AD classification. CNN models were trained on MRI data, with Softmax, SVM, and Random Forest classifiers used for evaluation, while the role of fully connected layers in enhancing classification performance was also examined.

A comprehensive survey in [10] reviewed over 80 studies applying ML to neuroimaging for AD detection. The findings revealed that SVM, decision trees, and ensemble approaches are the most frequently employed. The review also highlighted open challenges and opportunities for future research, such as applying metaheuristic optimization techniques like whale optimization and gray wolf optimization to improve feature selection.

In [11], the authors examined the robustness and discriminative capacity of various feature selection approaches for AD diagnosis. Their pipeline combined supervised and unsupervised methods within a bootstrap sampling-based framework, evaluated using the

ADNI dataset. Results indicated that balancing stability with discriminability significantly improves diagnostic performance. Similarly, [12] utilized data from the China Longitudinal Aging Study (CLAS) to build an interpretable predictive model for AD and MCI. By integrating five feature selection strategies with nine ML classifiers, the model achieved high sensitivity, specificity, and overall accuracy. Furthermore, SHapley Additive Explanations (SHAP) was applied to interpret the influence of individual features at both global and patient-specific levels, enhancing the clinical transparency of the system.

Finally, [13] explored multiple ML classifiers including decision trees, SVM, and neural networks, combined with feature selection methods such as Information Gain, Information Gain Ratio, Gini Index, Chi-Square, and PCA, using the OASIS dataset. Their experiments demonstrated that selecting the most informative features allowed classifiers to reach up to 90% accuracy, underscoring the importance of feature optimization in AD classification.

3. METHODOLOGY

3.1 Dataset

For this study, a total of 1,026 three-dimensional brain MRI scans were utilized, comprising an equal distribution of 402 images each for Cognitive Normal (CN), Alzheimer's Disease (AD), and Mild Cognitive Impairment (MCI). These T1-weighted structural MRI scans (1.5T) were obtained from the ADNI database (<http://adni.loni.usc.edu/>) [14]. Each scan had a slice thickness of 1.2 mm, with image dimensions of $256 \times 256 \times 166$, and voxel resolution of $1.0 \times 1.0 \times 1.2 \text{ mm}^3$.

3.2 Preprocessing

The MRI scans obtained from the ADNI database were preprocessed during acquisition to correct for Gradwarp, B1 non-uniformity, and N3 bias field distortions. Following this, N3 bias-corrected images were subjected to skull stripping using the deepbrain library in Python [16]. The resulting skull-stripped images were then linearly aligned to the 1 mm MNI152 template space using the FSL-FLIRT tool [17].

3.3 Feature Extraction

Feature extraction plays a vital role in highlighting the most informative properties of medical images, which is particularly important for disease classification tasks. In this study, we employed the **GLCM** method to derive texture-based features from MRI scans.

GLCM works by examining the spatial relationship between pixels within an image, quantifying how often pairs of pixels with specific intensity values occur in a given spatial configuration. Through this analysis, it generates a set of texture descriptors such as energy, homogeneity, contrast, entropy, correlation, and dissimilarity. These features provide valuable information about the underlying texture, including aspects like structural uniformity, similarity, intensity variations, randomness, pixel dependency, and localized changes in image patterns [18,19].

3.4 Feature Selection

Feature selection is a critical step in machine learning (ML), as it reduces dimensionality, improves computational efficiency, enhances model accuracy, and increases interpretability. In this study, we explored several feature selection strategies, including statistical tests, mutual information measures, dimensionality reduction methods, and embedded ML models such as tree-based classifiers and ensemble approaches.

1. **Analysis of Variance (ANOVA):** A statistical approach used to detect significant differences between group means. Within feature selection, ANOVA evaluates the discriminative power of features by comparing the variance across different groups to the variance within groups. Features with higher F-statistics are considered more influential [22].
2. **Chi-Square Test:** A method to evaluate the dependency between categorical features and the target class. Features yielding high Chi-Square scores indicate stronger associations with the target variable, making them more valuable for classification tasks [23].
3. **Mutual Information Classifier (MICIF):** Measures the degree of mutual dependence between attributes and the target. Features with higher mutual information values are identified as more predictive and relevant for classification [24].
4. **Principal Component Analysis (PCA):** A dimensionality reduction technique that projects high-dimensional data into a lower-dimensional space while retaining maximum variance. It produces new uncorrelated variables, called principal components, that capture the essential structure of the dataset [25].
5. **K-Nearest Neighbors (KNN):** A non-parametric algorithm that can be used for feature evaluation by analyzing how effectively features contribute to classification accuracy based on neighborhood similarity [26].
6. **Decision Tree Classifier (DTC):** Selects features by splitting nodes according to criteria such as Gini impurity or information gain. Features chosen at higher levels in the tree are typically more significant [25].
7. **Extra Trees Classifier (ETC):** An ensemble method that constructs multiple randomized decision trees. Feature importance is determined by the average impurity reduction achieved across all trees [27].
8. **Random Forest Classifier (RFC):** A bagging-based ensemble algorithm that aggregates multiple decision trees. Feature importance is computed by averaging the decrease in impurity contributed by each feature across the forest [28].
9. **Extreme Gradient Boosting (XGBoost, XGB):** A gradient boosting framework known for its efficiency and performance. It assesses feature importance using metrics such as gain (improvement brought by a feature), cover (frequency of usage in splits), and weight (number of times a feature is used) [29].

By leveraging this diverse set of feature selection methods, our objective is to isolate the most informative features from the dataset, thereby boosting the classification performance of Alzheimer's Disease (AD) stage prediction models.

3.5 Machine Learning Algorithms

To evaluate the classification task, several ML algorithms were employed, each offering unique strengths. The **K-Nearest Neighbors (KNN)** algorithm classifies samples based on the majority label among their k closest neighbors, making it an intuitive similarity-based method [26]. The **Decision Tree Classifier (DTC)**, on the other hand, hierarchically partitions data using feature-based splits, where internal nodes represent attribute tests and leaves indicate class outcomes [30]. Building on this, the **RFC** aggregates multiple decision trees to improve prediction robustness and reduce overfitting [28]. Ensemble boosting approaches, including the **Gradient Boosting Classifier (GBC)** [31] and **Adaptive Boosting**

(AdaBoost) [32], sequentially construct weak learners, with each new tree correcting the errors of previous models. The **CatBoost Classifier (CBC)** further enhances gradient boosting by efficiently managing categorical features and delivering high-performance results [33]. Probabilistic methods such as **Gaussian Naive Bayes (GNB)** assume independence between features and apply Bayes' theorem to make predictions efficiently [34]. The **Support Vector Machine (SVM)** identifies optimal hyperplanes in a high-dimensional space to maximize class separation [35], while the **Multi-Layer Perceptron (MLP)**, a type of feedforward artificial neural network, captures complex, non-linear patterns through its multiple layers of interconnected neurons [36].

4. RESULTS AND DISCUSSION

4.1 Evaluation Metrics

Multiple evaluation metrics were employed to assess the performance of the classification models. Accuracy represents the ratio of correctly predicted instances to the total number of predictions, providing an overall measure of performance. The ROC-AUC score reflects the model's ability to differentiate between classes, with the area under the ROC curve serving as an indicator of classification quality. Precision measures the proportion of correctly identified positive cases out of all predicted positives, emphasizing reliability in detecting true positives. Recall (Sensitivity) calculates the fraction of correctly predicted positive cases among all actual positives, showing how effectively the model captures relevant instances. Finally, the F1-Score, defined as the harmonic mean of precision and recall, offers a balanced metric that accounts for both false positives and false negatives, giving a more comprehensive performance evaluation.

4.2 Performance Comparison

From the comparative analysis, several important observations can be made. **Tree-based models**, including RFC and ETC, consistently outperformed other methods across most evaluation criteria. Notably, ETC achieved the best overall results highest accuracy, ROC-AUC, precision, recall, and F1-score, particularly when combined with PCA and its own feature selection mechanism. **Boosting algorithms**, such as GBC and XGB, also delivered strong results, with XGB standing out for its superior ROC-AUC performance, reflecting its ability to effectively separate different classes. Furthermore, the study shows that **feature selection techniques** like ANOVA and MICIF significantly improved the performance of complex classifiers such as GBC and XGB. Selecting the top 50% of features proved effective, striking a balance between dimensionality reduction and preserving essential discriminative information.

Selecting an appropriate combination of ML algorithms and feature engineering approaches is essential for maximizing classification accuracy. In this section, we present a comparative evaluation of various ML models on the AD dataset, integrating results obtained from different feature selection methods. The performance of the models was measured using standard evaluation metrics, including accuracy, ROC-AUC, precision, recall, and F1-score. Table I reports these outcomes across the employed feature selection techniques, emphasizing the influence of feature reduction on the classifiers' ability to differentiate between AD stages.

5. CONCLUSION

This research highlights the effectiveness of different ML algorithms and feature selection techniques in classifying AD stages. Ensemble and tree-based approaches such as Random Forest, Gradient Boosting, and XGBoost consistently outperformed other methods, achieving

higher accuracy, ROC-AUC, precision, recall, and F1-scores. The study also emphasizes the critical role of feature selection, with PCA and MICIF contributing notable improvements in performance. However, certain limitations should be acknowledged. The dataset may not comprehensively capture the full variability of AD progression, and incorporating additional clinical or demographic variables could enhance predictive power. Furthermore, applying advanced hyperparameter tuning strategies could yield further optimization of the models. For future work, exploring deep learning architectures, particularly CNNs, may provide deeper insights by capturing complex structural patterns present in neuroimaging data.

Table 1. Result of different ML algorithms based on various feature selection methods

Feature selection method	ANOVA										
Model	<i>KN N</i>	<i>DT C</i>	<i>GN B</i>	<i>SV M</i>	<i>ML P</i>	<i>RF C</i>	<i>GB C</i>	<i>XG B</i>	<i>AD B</i>	<i>ET C</i>	<i>CBC</i>
Accuracy	45. 62	56. 68	42. 39	56. 88	42. 39	56. 68	55. 76	62. 21	39. 63	59. 90	52.07
ROC-AUC Score	66. 31	67. 44	59. 66	73. 94	59. 27	75. 58	72. 45	77. 98	61. 94	77. 40	69.38
Precision	46. 78	56. 73	41. 85	56. 87	42. 31	56. 50	55. 48	62. 47	40. 50	59. 74	52.13
Recall	45. 47	56. 59	42. 06	56. 83	42. 26	56. 61	55. 64	62. 16	39. 57	59. 81	52.04
F1 Score	45. 18	56. 65	39. 13	56. 75	41. 99	56. 41	55. 52	62. 13	39. 79	59. 74	52.07
Mean Scores	50. 97	59. 38	45. 76	61. 11	46. 49	61. 29	59. 80	66. 20	45. 46	64. 20	56.41

Feature selection method	PCA										
Model	<i>KN N</i>	<i>DT C</i>	<i>GN B</i>	<i>SV M</i>	<i>ML P</i>	<i>RF C</i>	<i>GB C</i>	<i>XG B</i>	<i>AD B</i>	<i>ET C</i>	<i>CB C</i>
Accuracy	44.23	51.61	46.08	56.88	47.92	64.05	61.29	63.13	45.16	78.80	50.23
ROC-AUC Score	61.96	63.66	67.42	70.90	65.11	81.65	78.86	83.71	67.05	89.49	67.20
Precision	45.27	51.52	46.34	57.08	47.64	64.11	61.34	62.88	45.22	78.91	50.29
Recall	44.16	51.55	45.91	56.83	47.81	63.99	61.27	63.01	45.22	78.69	50.15
F1 Score	43.63	51.32	45.45	56.65	47.68	63.96	61.18	62.89	45.09	78.65	50.21
Mean Scores	48.77	54.53	51.32	60.38	52.09	68.44	65.67	68.15	50.63	81.46	54.48

Feature selection method	ETC										
Model	<i>KN</i> <i>N</i>	<i>DT</i> <i>C</i>	<i>GN</i> <i>B</i>	<i>SV</i> <i>M</i>	<i>ML</i> <i>P</i>	<i>RF</i> <i>C</i>	<i>GB</i> <i>C</i>	<i>XG</i> <i>B</i>	<i>AD</i> <i>B</i>	<i>ET</i> <i>C</i>	<i>CB</i> <i>C</i>
Accuracy	41. 94	53. 92	45. 62	51. 38	43. 78	59. 91	52. 53	58. 53	49. 31	61. 75	48. 85
ROC-AUC Score	60. 75	65. 51	60. 68	67. 19	60. 90	76. 10	70. 86	76. 77	65. 01	80. 63	66. 11
Precision	42. 86	54. 26	46. 64	51. 44	43. 22	59. 83	52. 60	58. 56	49. 96	61. 74	48. 40
Recall	41. 87	54. 02	45. 35	51. 33	43. 62	59. 78	52. 48	58. 41	49. 25	61. 64	48. 69
F1 Score	41. 38	53. 81	43. 20	51. 30	43. 31	59. 65	52. 49	58. 26	49. 33	61. 54	48. 33
Mean Scores	46. 73	56. 87	49. 04	55. 33	47. 80	63. 87	57. 12	63. 03	53. 40	66. 42	52. 92

Feature selection method	XGB										
Model	<i>KN</i> <i>N</i>	<i>DT</i> <i>C</i>	<i>GN</i> <i>B</i>	<i>SV</i> <i>M</i>	<i>ML</i> <i>P</i>	<i>RF</i> <i>C</i>	<i>GB</i> <i>C</i>	<i>XG</i> <i>B</i>	<i>AD</i> <i>B</i>	<i>ET</i> <i>C</i>	<i>CB</i> <i>C</i>
Accuracy	42. 86	49. 31	40. 09	50. 46	44. 24	53. 92	53. 92	59. 91	43. 78	56. 68	47. 93
ROC-AUC Score	60. 93	62. 01	58. 53	63. 73	62. 57	74. 26	72. 44	76. 44	65. 15	77. 77	66. 61
Precision	44. 58	50. 03	41. 52	53. 70	44. 18	53. 90	53. 83	60. 26	44. 49	56. 77	48. 31
Recall	42. 86	49. 35	39. 86	50. 35	44. 15	53. 90	53. 84	59. 80	43. 75	56. 56	47. 84
F1 Score	42. 95	49. 37	38. 36	49. 73	44. 09	53. 88	53. 83	59. 47	43. 88	56. 36	47. 77
Mean Scores	47. 83	52. 68	44. 62	54. 40	48. 77	58. 99	58. 50	64. 02	49. 32	61. 89	52. 65

ACKNOWLEDGMENT

We would like to acknowledge the ADNI repository for providing the neuroimaging data used in this study.

REFERENCES

1. Basheera, S., & Ram, M. S. S. (2019). Convolution neural network-based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 5, 974-986.

2. Angkoso, C. V., Tjahyaningtijas, H. P. A., Adrianto, Y., Sensusiaty, A. D., Purnama, I. K. E., & Purnomo, M. H. (2022). Multi-features fusion in multi-plane MRI images for Alzheimer's disease classification. *Int. J. Intell. Eng. Syst*, 15(4), 182-197.
3. Ammal, S. M., & Manoharan, P. S. (2023). Multi-Headed Deep Learning Models to Detect Abnormality of Alzheimer's Patients. *Computer Systems Science & Engineering*, 44(1).
4. AV, A., KUMAR, D. A. S., & LATIP, D. R. (2023). CNN-Mobilenetv2-Deep Learning-Based Alzheimer's Disease Prediction and Classification. *Journal of Theoretical and Applied Information Technology*, 101(9).
5. Lee, J., Meijer, E., Langa, K.M., Ganguli, M., Varghese, M., Banerjee, J., Khobragade, P., Angrisani, M., Kurup, R., Chakrabarti, S.S. and Gambhir, I.S. (2023). Prevalence of dementia in India: National and state estimates from a nationwide study. *Alzheimer's & Dementia*, 19(7), 2898-2912. <https://doi.org/10.1002/alz.12928>.
6. AlSaeed D, Omar SF.(2022) : Brain MRI Analysis for Alzheimer's Disease Diagnosis Using CNN-Based Feature Extraction and Machine Learning. *Sensors (Basel)*. 2022 Apr 11;22(8):2911. doi: 10.3390/s22082911. PMID: 35458896; PMCID: PMC9025443.
7. Barburiceanu, S.; Terebes, R.; Meza, S. (2021). 3D Texture Feature Extraction and Classification Using GLCM and LBP-Based Descriptors. *Appl. Sci.*, 11, 2332. <https://doi.org/10.3390/app11052332>.
8. Reddy, G & K, Nagireddy.(2021): A Robust Machine Learning Approach for Multiclass Alzheimer's Disease Detection using 3D Brain Magnetic Resonance Images. *Journal of Engineering Research*. 10.36909/jer.10511.
9. S. Loussaief and A. Abdelkrim, (2016). "Machine learning framework for image classification," *2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, Hammamet, Tunisia, pp. 58-61, doi: 10.1109/SETIT.2016.7939841.
10. Dara, O.A.; Lopez-Guede, J.M.; Raheem, H.I.; Rahebi, J.; Zulueta, E.; Fernandez-Gamiz, U.(2023) . Alzheimer's Disease Diagnosis Using Machine Learning: A Survey. *Appl. Sci*, 13, 8298. <https://doi.org/10.3390/app13148298>.
11. Gu F, Ma S, Wang X, Zhao J, Yu Y, Song X, (2022). Evaluation of Feature Selection for Alzheimer's Disease Diagnosis. *Front Aging Neurosci*;14:924113. doi: 10.3389/fnagi.2022.924113. PMID: 35813964; PMCID: PMC9263380.
12. Yue L, Chen W-g, Liu S-c, Chen S-b and Xiao S-f.(2023). An explainable machine learning based prediction model for Alzheimer's disease in China longitudinal aging study. *Frontiers in aging neuroscience* 15:1267020. doi: 10.3389/fnagi.2023.1267020.
13. Arjaria, S. K., Rathore, A. S., Bisen, D., & Bhattacharyya, S. (2024). Performances of machine learning models for diagnosis of Alzheimer's disease. *Annals of Data Science*, 11(1), 307-335.
14. Petersen, R.C.; Aisen, P.S.; Beckett, L.A.; Donohue, M.C.; Gamst, A.C.; Harvey, D.J.; Jack, C.R.; Jagust, W.J.; Shaw, L.M.; Toga, A.W. (2010); Alzheimer's disease neuroimaging initiative (ADNI): *Clinical characterization. Neurology*, 74, 201–209.

15. Altaf, T., Anwar, S. M., Gul, N., Majeed, M. N., Majid, M. (2018): Multi-class Alzheimer's Disease Classification Using Image and Clinical Features. *Biomedical Signal Processing and Control*, 43, 64-74. <https://doi.org/10.1016/j.bspc.2018.02.019>.
16. <https://github.com/SonicStrain/skull-stripping-using-deepbrain/tree/main>, accessed on 3/03/2023.
17. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012). Fsl. *NeuroImage*;62:782–790. doi: 10.1016/j.neuroimage.2011.09.015.
18. Haralick, R. M., Shanmugam, K., & Dinstein, I.(1973) . Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, ,SMC-3(6), 610-621.
19. Soh, L. K., & Tsatsoulis, C.(1999) . Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, ,37(2), 780-795.
20. Gonzalez, R. C., & Woods, R. E.(2002). Digital Image Processing (2nd ed.). *Prentice-Hall, Inc*.
21. Amadasun, M., & King, R. (1989). Textural features corresponding to textural properties. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(5), 1264-1274.
22. Montgomery, D. C.(2012). Design and Analysis of Experiments (8th ed.). *Hoboken, NJ: John Wiley & Sons*.
23. Agresti, A.(2007). An Introduction to Categorical Data Analysis (2nd ed.). *Hoboken, NJ: John Wiley & Sons*.
24. Brown, G., Pocock, A., Zhao, M.J., & Luján, M. (2012). Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research*, 13, 27-66.
25. Rupapara, V., Rustam, F., Ishaq, A., Lee, E., Ashraf, I. (2023). Chi-square and PCA based feature selection for diabetes detection with ensemble classifier. *Intelligent Automation & Soft Computing*, 36(2), 1931-1949.
26. Altman, N.S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*, 46(3), 175-185.
27. Geurts, P., Ernst, D., & Wehenkel, L.(2006) . Extremely randomized trees. *Machine Learning*, 63(1), 3-42.
28. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
29. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
30. Ranjit Panigrahi, Samarjeet Borah, (2019). "Classification and Analysis of Facebook Metrics Dataset Using Supervised Classifiers", *Social Network Analytics*, Academic Press , Pages 1-19,ISBN 9780128154588,<https://doi.org/10.1016/B978-0-12-815458-8.00001-3>.
31. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232.

32. Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139.
33. Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.
34. Hand, D. J., & Yu, K. (2001). Idiot's Bayes—not so stupid after all? *International Statistical Review*, 69(3), 385-398.
35. Cortes, C., & Vapnik, V. (1995) Support-vector networks. *Machine Learning*, 20(3), 273-297.
36. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.