# Comparative Evaluation of Twenty Machine Learning Classifiers for Binary Dementia Detection Using Longitudinal Clinical and MRI-Derived Volumetric Features

**Priyanka Kaushik[1]*, Dr. Anand Singh Bisen[2]**

1 Research Scholar, School of Engineering & Technology, Gwalior, Madhya Pradesh, India

kaushik.priyanka17@gmail.com

2 Director, Engineering, Vikrant University, Gwalior, Madhya Pradesh, India

**Abstract:** Alzheimer's disease (AD) accounts for the majority of the more than 58 million dementia cases recorded worldwide, and no disease-modifying therapy is currently available. Because intervention is most effective before irreversible neuronal loss occurs, low-cost screening instruments that operate on routinely collected clinical variables are of considerable practical value. This study reports a systematic comparison of twenty supervised machine learning classifiers for binary discrimination between demented and non-demented participants using the longitudinal Open Access Series of Imaging Studies (OASIS) cohort. Nine attributes were used: age, gender, years of education, socioeconomic status, Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), estimated total intracranial volume (eTIV), normalised whole-brain volume (nWBV) and atlas scaling factor (ASF). Participants originally labelled "Converted" were merged into the demented class to yield an unambiguous two-class target. Missing continuous values were median-imputed, categorical values mode-imputed, all continuous attributes standardised to zero mean and unit variance, and principal component analysis retained the components accounting for 95% of the variance. Models were assessed under an 80:20 split with five-fold cross-validation using precision, recall, F1-score and accuracy. Tree-based ensembles dominated the ranking: the Extra Trees classifier achieved the best overall balance at 87% accuracy with 91% precision and 85% recall on the demented class, followed by hard voting (85%) and XGBoost (84%). Scaling-sensitive and independence-assuming learners — stochastic gradient descent, Bernoulli and multinomial naïve Bayes, and bagging — performed poorest, between 56% and 63%. Correlation analysis identified MMSE, CDR and nWBV as the dominant discriminative attributes. The results establish a reproducible tabular baseline against which multimodal imaging-plus-clinical architectures can be judged, and indicate that structured cognitive and volumetric measures alone support useful, but not clinically sufficient, dementia screening.

**Keywords:** Alzheimer's disease; dementia screening; machine learning; ensemble classifiers; Extra Trees; OASIS; Mini-Mental State Examination; Clinical Dementia Rating; brain atrophy; class imbalance.

## 1. INTRODUCTION

Alzheimer's disease is a progressive neurodegenerative disorder and the leading cause of dementia worldwide. It erodes memory, reasoning and executive function until affected individuals lose the capacity to live independently. Susceptibility rises sharply with age and is

estimated to approximately double with every five-year increment beyond 65 years. The World Health Organization reports that more than 58 million people currently live with Alzheimer's disease or a related dementia, and projects that this population will reach 154 million by 2050. Alzheimer's disease alone represents roughly 65–75% of all dementia cases, which places it among the most consequential public-health problems of the coming decades.

Biologically, the disorder is defined by two pathological hallmarks: extracellular amyloid-beta plaques and intraneuronal neurofibrillary tangles composed of hyperphosphorylated tau protein. Their accumulation disrupts synaptic transmission, drives neuronal death and produces measurable cerebral atrophy, most prominently in hippocampal and cortical regions. Despite decades of investigation, no curative therapy exists; available pharmacological agents such as donepezil, rivastigmine and memantine offer symptomatic relief without arresting the underlying degeneration. Early identification therefore remains the principal lever for improving outcomes, because it permits therapeutic and lifestyle measures to begin while cognitive reserve is still substantial.

Conventional diagnosis combines clinical examination, neuropsychological testing and neuroimaging. Instruments such as the MMSE and the CDR quantify cognitive status, while magnetic resonance imaging (MRI), computed tomography and positron emission tomography reveal structural and functional abnormalities. Each modality carries constraints. Clinical judgement varies between assessors and can miss the subtle changes characteristic of the earliest disease stage. Imaging is more objective but expensive, resource-intensive and frequently unavailable in low- and middle-income settings; interpreting neuroimaging also demands specialist expertise that does not scale to population screening.

Against this background, machine learning applied to structured clinical records offers an attractive intermediate option. Cognitive scores, demographic attributes and volumetric summaries derived from a single structural scan are inexpensive to obtain, are already recorded in most memory-clinic workflows, and can be modelled with algorithms that are computationally light enough to run on commodity hardware. The open question is not whether such models can be fitted, but which family of algorithms extracts the most reliable signal from a small, imbalanced and partially incomplete tabular cohort, and how far that signal can be pushed before imaging features become indispensable.

This paper addresses that question through a controlled comparison of twenty classifiers spanning tree-based ensembles, linear and discriminant models, probabilistic learners, instance-based methods, a feed-forward neural network and two voting ensembles, all evaluated on an identical preprocessing pipeline and identical cross-validation protocol. The contribution is threefold: a uniformly reported benchmark across an unusually broad algorithm set; an analysis of why particular families succeed or fail on this data type; and a quantified performance ceiling for clinical-only dementia screening that motivates subsequent multimodal work.

## 2. RELATED WORK

Machine learning applied to Alzheimer's prediction has developed along three broad lines. The first uses neuropsychological and demographic variables directly. Kavitha et al. demonstrated early-stage prediction from cognitive scores and reported that ensemble tree methods outperformed single learners on the OASIS cohort. Joshi et al. classified disease stages using several conventional algorithms and established the value of MMSE and CDR as primary discriminators, a finding that recurs consistently across the subsequent literature. Vidushi and Shrivastava, and Vijayalakshmi and Sabharwal, reached comparable conclusions with different algorithm selections, though neither study reported a systematic comparison across model families.

A second line derives quantitative descriptors from MRI and treats them as tabular inputs. Khan and Zubair applied dimensionality reduction to cross-sectional MRI features and showed that a compact representation preserves most of the discriminative information, and later extended this to longitudinal exploratory analysis. Battineni et al. combined imaging-derived measures with clinical variables using multimodal machine learning and reported improved detection relative to either source alone. Gopagoni et al. and Kishore et al. evaluated conventional algorithms on comparable feature sets, again finding ensemble methods advantageous.

A third line explores alternative biomarkers. Eke et al. used support vector machines on blood plasma proteins; Peña-Bautista et al. examined lipid peroxidation markers; Ammar and Ayed applied speech processing; Yang et al. reviewed M/EEG-derived markers; and Saied et al. investigated radio-frequency signals. These modalities are promising but remain distant from routine deployment.

Two limitations recur across this body of work. First, studies typically report a small number of algorithms, which makes cross-paper comparison unreliable because preprocessing, splits and metrics differ. Buyrukoğlu addressed part of this by studying ensemble learning explicitly, but the comparison set remained narrow. Second, class imbalance is frequently acknowledged and rarely quantified in its effect on recall for the minority class — which is precisely the clinically important quantity, since a missed case of dementia is more costly than a false alarm. The present study responds to both limitations by holding the pipeline fixed across twenty algorithms and by reporting per-class precision and recall rather than accuracy alone.

## 3. MATERIALS AND METHODS

### *A. Dataset*

Experiments used the longitudinal Open Access Series of Imaging Studies (OASIS) cohort, which records repeated observations of the same individuals and therefore captures both structural change and cognitive decline over time. Each participant record combines demographic attributes, neuropsychological scores and volumetric summaries derived from structural MRI acquired under standardised protocols. The attributes used are listed in Table I.

**Table 1: Clinical and volumetric attributes used for classification**

| Attribute | Description and diagnostic role |
|---|---|
| Age | Participant age in years; the dominant demographic risk factor for dementia. |
| Gender | Male/female indicator, reflecting demographic differences in prevalence and progression. |
| Education | Years of formal education; a proxy for cognitive reserve, which may delay symptom onset. |
| SES | Socioeconomic status, a composite of income, education and occupation capturing lifestyle and environmental risk. |

| Attribute | Description and diagnostic role |
|---|---|
| MMSE | Mini-Mental State Examination, scored 0–30; quantifies severity of cognitive impairment. |
| CDR | Clinical Dementia Rating, scored 0–3; the standard staging scale and the basis of the class label. |
| eTIV | Estimated total intracranial volume; used to normalise regional volumes for head size. |
| nWBV | Normalised whole-brain volume; the ratio of brain tissue to intracranial volume, indexing atrophy. |
| ASF | Atlas scaling factor from registration to a standard template, supporting cross-participant volume normalisation. |

Participants were originally distributed across three groups — Demented, Non-Demented and Converted. The Converted group comprises individuals who presented as non-demented at baseline and subsequently progressed. Because an intermediate label of this kind is ambiguous for a two-class formulation, converted participants were merged into the demented class, yielding targets of Demented (1) and Non-Demented (0). This is clinically defensible: a participant who has progressed to mild or early dementia belongs on the disease side of a screening decision. Exploratory inspection nevertheless confirmed a clear residual imbalance in favour of the non-demented class, which is consistent with population prevalence and which motivated the use of per-class metrics throughout.

### *B. Preprocessing*

Real clinical records are routinely incomplete because individual tests are skipped or demographic fields are left unrecorded. Continuous attributes with missing entries — principally MMSE — were imputed with the feature median, and categorical attributes with the feature mode, so that every participant contributed a complete vector and no record was discarded. All continuous attributes were then standardised to zero mean and unit variance,

which prevents variables measured on wide numeric ranges, such as eTIV, from dominating distance- and gradient-based learners and improves optimisation stability.

Principal component analysis was applied to the standardised matrix, retaining the components that jointly explain 95% of the variance. Several of the volumetric attributes are strongly interdependent by construction — ASF is derived from the same registration that produces eTIV — so this step removes redundancy while preserving nearly all discriminative information, and reduces the dimensionality that distance-based and probabilistic learners must contend with.

### *C. Classifier Families*

Twenty algorithms were evaluated, grouped into six families so that observed behaviour can be attributed to modelling assumptions rather than to individual implementations:

- **Tree-based methods:** Decision Tree, Random Forest, Extra Trees, Gradient Boosting, Histogram-based Gradient Boosting, XGBoost, AdaBoost and Bagging. These partition the feature space hierarchically, tolerate mixed attribute types and outliers, and expose feature-importance scores that support clinical interpretation.
- **Linear and discriminant models:** Logistic Regression, Linear Discriminant Analysis, Quadratic Discriminant Analysis and a Stochastic Gradient Descent classifier. These assume linear or class-conditional Gaussian structure and serve as interpretable, fast baselines.
- **Probabilistic models:** Gaussian, Bernoulli and Multinomial naïve Bayes, which assume conditional independence between attributes and provide calibrated probability estimates cheaply.
- **Instance-based:** k-Nearest Neighbours, which classifies by local similarity and requires no explicit training phase.
- **Neural-network-inspired:** a Multi-Layer Perceptron with nonlinear hidden activations, capable of representing higher-order attribute interactions and acting as a bridge to deep architectures.
- **Hybrid ensembles:** hard voting, which takes the majority predicted class, and soft voting, which averages predicted probabilities across constituent learners.

### *D. Experimental Protocol and Metrics*

The preprocessed data were divided 80:20 into training and test partitions, and five-fold cross-validation was applied within the training partition to obtain a stable estimate of generalisation. Every algorithm received identical folds and identical preprocessing so that differences in outcome reflect the learner rather than the pipeline. Performance was reported as precision, recall, F1-score and support for each class separately, together with overall accuracy. Per-class reporting is essential here: on an imbalanced screening task, an algorithm can attain deceptively high accuracy by favouring the majority class while failing on precisely the cases that matter. Receiver-operating-characteristic curves and confusion matrices were additionally examined to assess threshold-independent discrimination and to characterise the direction of residual errors.

## 4. RESULTS AND DISCUSSION

### *A. Exploratory Findings*

Exploratory analysis preceded model fitting. Violin plots of CDR by gender indicated slightly greater variability in mild-to-moderate ratings among female participants, while CDR plotted against age showed both a rising central tendency and broader dispersion in older participants — consistent with the clinical understanding of Alzheimer's disease as an age-associated disorder whose severity becomes increasingly heterogeneous after 65 years.

Pairwise distribution analysis produced two physiologically coherent relationships that proved central to model behaviour. Age correlated negatively with normalised whole-brain volume, reflecting progressive atrophy, and MMSE correlated positively with nWBV, indicating that preserved cognition accompanies preserved brain tissue. The correlation heatmap further showed a strong negative association between MMSE and CDR, confirming that these two instruments measure the same underlying construct from opposite directions. These findings guided both feature retention and the interpretation of the importance rankings produced by the tree-based models.

### *B. Comparative Classifier Performance*

Table II reports the complete comparison. Ensemble tree methods occupy the upper region of the ranking without exception. The Extra Trees classifier produced the strongest and most balanced result at 87% accuracy, with 91% precision and 85% recall on the demented class

and 84% precision and 90% recall on the non-demented class. Hard voting and XGBoost followed at 85% and 84% accuracy respectively, with soft voting and Random Forest close behind.

**Table 2: Per-class performance of twenty classifiers for binary dementia detection (%)**

| # | Classifier | Class | Prec. | Rec. | F1 | Supp. | Acc. |
|---|---|---|---|---|---|---|---|
| 1 | Extra Trees (ET) | Non-demented (0) | 84 | 90 | 87 | 52 | 87 |
| | | Demented (1) | 91 | 85 | 88 | 60 | |
| 2 | Hard Voting (HV) | Non-demented (0) | 82 | 87 | 84 | 52 | 85 |
| | | Demented (1) | 88 | 83 | 85 | 60 | |
| 3 | Soft Voting (SV) | Non-demented (0) | 81 | 85 | 83 | 52 | 84 |
| | | Demented (1) | 86 | 83 | 85 | 60 | |
| 4 | XGBoost (XGB) | Non-demented (0) | 81 | 85 | 83 | 52 | 84 |
| | | Demented (1) | 86 | 83 | 85 | 60 | |
| 5 | Random Forest (RF) | Non-demented (0) | 80 | 87 | 83 | 52 | 83 |
| | | Demented (1) | 88 | 82 | 84 | 60 | |
| 6 | AdaBoost (ADB) | Non-demented (0) | 78 | 87 | 82 | 52 | 82 |
| | | Demented (1) | 87 | 78 | 82 | 60 | |

| # | Classifier | Class | Prec. | Rec. | F1 | Supp. | Acc. |
|---|---|---|---|---|---|---|---|
| 7 | Decision Tree (DT) | Non-demented (0) | 75 | 83 | 79 | 52 | 79 |
| | | Demented (1) | 84 | 77 | 80 | 60 | |
| 8 | Quadratic Discriminant (QDA) | Non-demented (0) | 72 | 90 | 80 | 52 | 79 |
| | | Demented (1) | 89 | 70 | 79 | 60 | |
| 9 | Hist. Gradient Boosting | Non-demented (0) | 75 | 85 | 79 | 52 | 79 |
| | | Demented (1) | 85 | 75 | 80 | 60 | |
| 10 | Multi-Layer Perceptron (MLP) | Non-demented (0) | 82 | 69 | 75 | 52 | 78 |
| | | Demented (1) | 76 | 87 | 81 | 60 | |
| 11 | Support Vector Machine (SVM) | Non-demented (0) | 71 | 88 | 79 | 52 | 78 |
| | | Demented (1) | 87 | 68 | 77 | 60 | |
| 12 | Gradient Boosting (GBC) | Non-demented (0) | 76 | 79 | 77 | 52 | 78 |
| | | Demented (1) | 81 | 78 | 80 | 60 | |
| 13 | Linear Discriminant (LDA) | Non-demented (0) | 68 | 88 | 77 | 52 | 75 |
| | | Demented (1) | 86 | 63 | 73 | 60 | |

| # | Classifier | Class | Prec. | Rec. | F1 | Supp. | Acc. |
|---|---|---|---|---|---|---|---|
| 14 | Logistic Regression (LR) | Non-demented (0) | 69 | 77 | 73 | 52 | 73 |
| | | Demented (1) | 78 | 70 | 74 | 60 | |
| 15 | Gaussian Naïve Bayes (GNB) | Non-demented (0) | 67 | 85 | 75 | 52 | 73 |
| | | Demented (1) | 83 | 63 | 72 | 60 | |
| 16 | k-Nearest Neighbours (KNN) | Non-demented (0) | 62 | 75 | 68 | 52 | 67 |
| | | Demented (1) | 73 | 60 | 66 | 60 | |
| 17 | SGD Classifier | Non-demented (0) | 56 | 98 | 71 | 52 | 63 |
| | | Demented (1) | 95 | 33 | 49 | 60 | |
| 18 | Multinomial Naïve Bayes (MNB) | Non-demented (0) | 58 | 67 | 63 | 52 | 62 |
| | | Demented (1) | 67 | 58 | 63 | 60 | |
| 19 | Bagging Classifier | Non-demented (0) | 56 | 69 | 62 | 52 | 61 |
| | | Demented (1) | 67 | 53 | 59 | 60 | |
| 20 | Bernoulli Naïve Bayes (BNB) | Non-demented (0) | 52 | 65 | 58 | 52 | 56 |

*Per-class values for the demented class of the Bernoulli naïve Bayes model were not recorded in the experiment log; overall accuracy is reported.*

### *C. Threshold-Independent Discrimination*

Receiver-operating-characteristic analysis reinforced the ranking obtained from point metrics. Extra Trees, XGBoost, Random Forest and the hard-voting ensemble produced curves closest to the upper-left corner, indicating superior separation across the full range of decision thresholds; Extra Trees attained the highest area under the curve, consistent with its advantage in precision and F1-score. XGBoost and hard voting recorded areas in the region of 0.90 to 0.92. Comparing training and testing curves for the ensemble models showed closely matched areas, indicating that the observed performance is not the product of overfitting to the training partition. Stochastic gradient descent and Bernoulli naïve Bayes produced markedly flatter curves, confirming limited discriminative capacity rather than a poorly chosen operating point.

Confusion matrices showed the same pattern in a different form. Random Forest, Extra Trees and XGBoost concentrated observations along the main diagonal, with Extra Trees producing the cleanest separation and the fewest misclassifications. Decision Tree and quadratic discriminant analysis exhibited more off-diagonal mass, which is consistent with overfitting on a small attribute set in the former case and with the strong distributional assumption in the latter.

### *D. Discussion*

The dominance of tree-based ensembles is explicable in terms of the data. Nine attributes, several of them interdependent and none of them linearly separable with respect to the target, present exactly the conditions under which recursive partitioning succeeds: the models capture nonlinear interactions such as the joint effect of low MMSE and reduced nWBV without requiring those interactions to be specified in advance, and aggregation across many trees suppresses the variance that a single deep tree would incur on a cohort of this size.

The failure modes are equally informative. The stochastic gradient descent classifier achieved 98% recall on the non-demented class but only 33% on the demented class — the signature of a model that has effectively collapsed toward the majority class, and a result that would be entirely concealed by its 63% overall accuracy. This illustrates why per-class reporting is not optional on imbalanced clinical tasks. The naïve Bayes variants suffered from their independence assumption, which is untenable when MMSE and CDR are strongly negatively correlated and eTIV, nWBV and ASF are analytically linked. k-Nearest Neighbours degraded

because Euclidean distance is a weak similarity measure in a standardised nine-dimensional space where attribute relevance is highly uneven.

The comparable performance of hard and soft voting relative to their strongest constituent members indicates that aggregation stabilises predictions but does not create information that the base learners lack. Extra Trees, which already randomises split thresholds aggressively, extracts most of the available signal on its own.

The practical implication is that a clinical-only screening model built on routinely available cognitive and volumetric summaries reaches approximately 87% accuracy with 85% sensitivity to the demented class. That level supports triage — identifying who should proceed to specialist assessment — but leaves roughly one in seven cases misclassified, which is insufficient for diagnostic use and does not address stage discrimination at all, since the formulation is binary. Both limitations point toward the incorporation of imaging representations, which encode spatial atrophy patterns that no scalar volumetric summary can capture.

### *E. Limitations*

Several constraints qualify these results. The cohort is drawn from a single public repository and may not reflect the demographic and clinical diversity of routine practice. Diagnostic labels derived from CDR are uncertain near stage boundaries, and merging the converted group into the demented class, while clinically reasonable, removes a genuinely intermediate category. Median and mode imputation preserves sample size but attenuates variance in the affected attributes. The evaluation is retrospective and reports no external validation cohort, so the reported figures should be read as a comparative ranking under a fixed protocol rather than as estimates of deployment performance.

## 5. CONCLUSION

This study compared twenty supervised classifiers for binary dementia detection under a single fixed preprocessing and evaluation protocol applied to longitudinal clinical and MRI-derived volumetric attributes. Tree-based ensembles consistently outperformed linear, probabilistic and instance-based alternatives, with the Extra Trees classifier reaching 87% accuracy, 91% precision and 85% recall on the demented class, ahead of hard voting at 85% and XGBoost at 84%. Scaling-sensitive and independence-assuming learners performed substantially worse,

and per-class analysis revealed that overall accuracy conceals severe minority-class failure in the weakest models. MMSE, CDR and normalised whole-brain volume emerged as the dominant discriminative attributes, in agreement with established clinical understanding.

The benchmark establishes both a reproducible reference point and a ceiling. Structured cognitive and volumetric data support useful screening but cannot represent the spatial patterns of cortical thinning and hippocampal atrophy that distinguish adjacent disease stages. Subsequent work will therefore extend this baseline toward multimodal architectures that combine these clinical attributes with convolutional representations learned directly from structural MRI, and will evaluate the resulting models on multiclass stage discrimination with explicit interpretability analysis.

**References**

1. World Health Organization, "Dementia," 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

2. Down To Earth, "Dementia cases in India to double by 2050, Lancet study shows," 2021.

3. C. Kavitha, V. Mani, S. R. Srividhya, O. I. Khalaf, and C. A. Tavera Romero, "Early-stage Alzheimer's disease prediction using machine learning models," Frontiers in Public Health, vol. 10, 853294, 2022.

4. S. Joshi, G. G. V. Simha, D. P. Shenoy, K. R. Venugopal, and L. M. Patnaik, "Classification and treatment of different stages of Alzheimer's disease using various machine learning methods," International Journal of Bioinformatics Research, vol. 2, no. 1, 2010.

5. S. Joshi, D. Shenoy, G. V. Simha, P. L. Rashmi, K. R. Venugopal, and L. M. Patnaik, "Classification of Alzheimer's disease and Parkinson's disease using machine learning and neural network methods," in Proc. 2nd Int. Conf. Machine Learning and Computing, 2010, pp. 218–222.

6. G. Battineni, M. A. Hossain, N. Chintalapudi, E. Traini, V. R. Dhulipalla, M. Ramasamy, and F. Amenta, "Improved Alzheimer's disease detection by MRI using multimodal machine learning algorithms," Diagnostics, vol. 11, no. 11, 2103, 2021.

7. A. Khan and S. Zubair, "A machine learning-based robust approach to identify dementia progression employing dimensionality reduction in cross-sectional MRI data," in Proc. IEEE Conf., 2020.

8. A. Khan and S. Zubair, "Longitudinal magnetic resonance imaging as a potential correlate in the diagnosis of Alzheimer disease: exploratory data analysis," JMIR Biomedical Engineering, vol. 5, no. 1, e14389, 2020.

9. D. R. Gopagoni, P. V. Lakshmi, and R. S. Ananthula, "An application of machine learning strategies to predict Alzheimer's illness progression in patients," International Journal of Advanced Research in Engineering and Technology, vol. 11, no. 6, pp. 1056–1063, 2020.

10. P. Kishore, C. U. Kumari, M. N. V. S. S. Kumar, et al., "Detection and analysis of Alzheimer's disease using various machine learning algorithms," Elsevier, 2020.

11. A. R. Vidushi and A. K. Shrivastava, "Diagnosis of Alzheimer disease using machine learning approaches," International Journal of Advanced Science and Technology, vol. 29, pp. 7062–7073, 2019.

12. S. Vijayalakshmi and M. Sabharwal, "Alzheimer's disease detection through machine learning," Annals of the Romanian Society for Cell Biology, vol. 25, no. 3, pp. 2782–2792, 2021.

13. K. M. M. Uddin, M. J. Alam, M. A. Uddin, and S. Aryal, "A novel approach utilizing machine learning for the early diagnosis of Alzheimer's disease," Biomedical Materials & Devices, vol. 1, no. 2, pp. 882–898, 2023.

14. S. Buyrukoğlu, "Improvement of machine learning models' performances based on ensemble learning for the detection of Alzheimer disease," in Proc. 6th Int. Conf. Computer Science and Engineering (UBMK), 2021, pp. 102–106.

15. C. S. Eke, E. Jammeh, X. Li, C. Carroll, S. Pearson, and E. Ifeachor, "Early detection of Alzheimer's disease with blood plasma proteins using support vector machines," IEEE J. Biomedical and Health Informatics, vol. 25, no. 1, pp. 218–226, 2020.

16. C. Peña-Bautista, C. Vigor, J. M. Galano, C. Oger, T. Durand, I. Ferrer, and C. Cháfer-Pericás, "Plasma lipid peroxidation biomarkers for early and non-invasive Alzheimer disease detection," Free Radical Biology and Medicine, vol. 124, pp. 388–394, 2018.

17. R. B. Ammar and Y. B. Ayed, "Speech processing for early Alzheimer disease diagnosis: machine learning based approach," in Proc. IEEE/ACS 15th Int. Conf. Computer Systems and Applications (AICCSA), 2018, pp. 1–8.

18. S. Yang, J. M. S. Bornot, K. Wong-Lin, and G. Prasad, "M/EEG-based biomarkers to predict MCI and Alzheimer's disease: a review from the machine learning perspective," IEEE Trans. Biomedical Engineering, vol. 66, no. 10, pp. 2924–2935, 2019.

19. I. M. Saied, T. Arslan, and S. Chandran, "Classification of Alzheimer's disease using RF signals and machine learning," IEEE J. Electromagnetics, RF and Microwaves in Medicine and Biology, vol. 6, no. 1, pp. 77–85, 2021.

20. C. H. Chang, C. H. Lin, and H. Y. Lane, "Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease," International Journal of Molecular Sciences, vol. 22, no. 5, 2761, 2021.

21. O. V. Forlenza, M. Radanovic, L. L. Talib, I. Aprahamian, B. S. Diniz, H. Zetterberg, and W. F. Gattaz, "Cerebrospinal fluid biomarkers in Alzheimer's disease: diagnostic accuracy and prediction of dementia," Alzheimer's & Dementia: DADM, vol. 1, no. 4, pp. 455–463, 2015.

22. D. G. Olle Olle, J. Zoobo Bisse, and G. Abessolo Alo'o, "Application and comparison of K-means and PCA based segmentation models for Alzheimer disease detection using MRI," Discover Artificial Intelligence, vol. 4, no. 1, 11, 2024.

23. S. Mohsen, "Alzheimer's disease detection using deep learning and machine learning: a review," Artificial Intelligence Review, vol. 58, 262, 2025.