International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04812
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

An Explainable Ensemble Machine Learning Framework for Early Detection of Alzheimer's Disease Using Clinical Data

Kanimozhi Rajkumar[1*],

Lalitha Kishore Jothi Ravikumar[1],

Al Mushavir Rahman Mohamed Iliyas[1],

Sowmiya Parisutha Packiyam[1],

Gopal Samy Balakrishnan[1]

[1] Department of Biotechnology, KIT - Kalaignarkarunanidhi Institute of Technology, Coimbatore, India

*Corresponding author: Kanimozhi Rajkumar,

(kanimozhi.rajkumar@gmail.com)

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia, accounting for about 60–80% of cases worldwide. It is characterised by a gradual decline in cognitive functions such as memory, thinking, orientation, comprehension, and communication. The underlying pathology involves the buildup of beta-amyloid plaques and tau protein tangles in the brain, which cause neuronal damage and brain shrinkage over time. Clinically, symptoms range from mild forgetfulness to severe cognitive impairment, eventually impairing daily functioning. Globally, over 55 million people have dementia, with Alzheimer's being the primary contributor. The societal and economic impacts are enormous, with dementia-related costs exceeding $1 trillion annually. A major challenge is that AD is often diagnosed only in later stages, after extensive neurodegeneration has occurred. Early detection is essential to slow progression, enhance quality of life, and optimise treatment options.

Keywords – Alzheimer's Disease (AD), Machine Learning (ML), Early Detection, Explainability, Ensemble Learning, Clinical Data, Diagnostic Framework.

1. Introduction

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder and the leading cause of **dementia**, accounting for approximately 60–80% of cases worldwide (World Health Organization, 2025). It is characterized by a gradual decline in cognitive functions such as memory, thinking, orientation, comprehension, and communication. The underlying pathology involves the buildup of **beta-amyloid plaques** and **tau protein tangles** in the brain, which cause neuronal damage and brain shrinkage over time. Clinically, symptoms range from mild forgetfulness to severe cognitive impairment, eventually impairing daily functioning.

Globally, over 55 million people have dementia, with Alzheimer's being the primary contributor. The societal and economic impacts are enormous, with dementia-related costs exceeding $1 trillion annually (Alzheimer's Disease International, 2023). A major challenge is that AD is often diagnosed only in later stages, after extensive neurodegeneration has occurred. **Early detection** is essential to slow progression, enhance quality of life, and optimize treatment options (Kavitha et al., 2022).

Traditional diagnostic methods including **magnetic resonance imaging (MRI)**, **positron emission tomography (PET)**, **cerebrospinal fluid (CSF) analysis**, and **neuropsychological testing** are informative but often invasive, expensive, or unavailable in low-resource settings (Illakiya & Karthik, 2023). These approaches also demand substantial clinical expertise and infrastructure. Conversely, the expanding field of **machine learning (ML)** presents a promising, scalable, and non-invasive alternative for early diagnosis by analyzing routine clinical and behavioral data (Kavitha et al., 2022).

In the last decade, ML models have been widely studied for their potential in early AD detection (Kavitha et al., 2022; Ortiz et al., 2016). However, many depend heavily on imaging or genetic data and often lack transparency, making them less suitable for clinical use. While **black-box models** may achieve high accuracy, they do not reveal the reasoning behind their predictions—a critical aspect for clinical acceptance (Hamoud et al., 2025; Linardatos et al., 2020).

This study introduces an **explainable ML approach** using structured clinical data. It evaluates five classifiers, including ensemble and deep learning models, emphasizing both predictive accuracy and interpretability through **SHAP**.

2. Related Work

Recent years have seen extensive efforts to apply **machine learning (ML)** for predicting and diagnosing Alzheimer's Disease (AD) (Kavitha et al., 2022; Sarma & Chatterjee, 2025). Much of this research has focused on **neuroimaging modalities** like **MRI** and **PET scans**, which offer detailed structural and functional insights into the brain. For example, Hosseini-Asl et al. (2016) used **convolutional autoencoders** to extract key features from MRI data, achieving competitive accuracy but requiring computationally intensive preprocessing. **Ensemble methods** such as **Random Forests** (Ortiz et al., 2016) and **Gradient Boosting** have shown strong classification performance in these imaging studies. Although **deep neural networks (DNNs)** are powerful, they often lack interpretability and demand large labelled datasets (Illakiya & Karthik, 2023). **Hybrid models** combining neuroimaging with biomarkers and clinical scores have improved diagnosis but remain less accessible due to their complexity and data costs (Ritter et al., 2015).

In parallel, researchers have explored structured clinical data, beyond imaging, for early detection (Kavitha et al., 2022; Nallapu et al., 2024). Kim et al. (2021) demonstrated the predictive value of demographic and behavioral features using models like **Support Vector Machines (SVM)**, **Decision Trees**, and **Naïve Bayes** applied to neuropsychological assessments. Studies have also examined modifiable risk factors—such as cardiovascular health, obesity, and social isolation—using indices like LIBRA and frameworks like In-MINDD.

Despite these advances, there's a notable lack of focus on **model transparency** (Hamoud et al., 2025; Linardatos et al., 2020). Many existing studies prioritize accuracy over explainability, which is crucial for clinical adoption (Linardatos et al., 2020; Prokhorenkova et al., 2018). Limited use has been made of **interpretability tools** like **SHAP** that can offer feature-level insights. This study addresses these gaps by developing a comprehensive, explainable ML framework based on non-imaging clinical data. Using SHAP for interpretation and ensemble modeling, it emphasizes both predictive performance and model transparency, ensuring better applicability in real-world scenarios.

3. Data and Methodology

3.1 Dataset

This study utilizes the Alzheimer's disease dataset (Kharoua, 2024) publicly hosted on Kaggle. The dataset comprises 15,000+ patient records containing structured clinical and behavioural variables, including age, gender, Mini-Mental State Examination (MMSE) scores, memory complaint frequency, Activities of Daily Living (ADL) status, cognitive functioning, and physician observations. The target variable is a binary diagnosis indicating whether a subject is classified as cognitively healthy or potentially at risk for Alzheimer's Disease.

The dataset was chosen for its rich variety of real-world attributes, its accessibility, and its relevance to non-invasive and scalable diagnostic tools. The balance between clinical and behavioural metrics makes it particularly suitable for training interpretable machine learning models without requiring expensive imaging modalities.

3.2 Preprocessing

Data preprocessing was a critical step in ensuring model robustness. The following operations were applied sequentially:

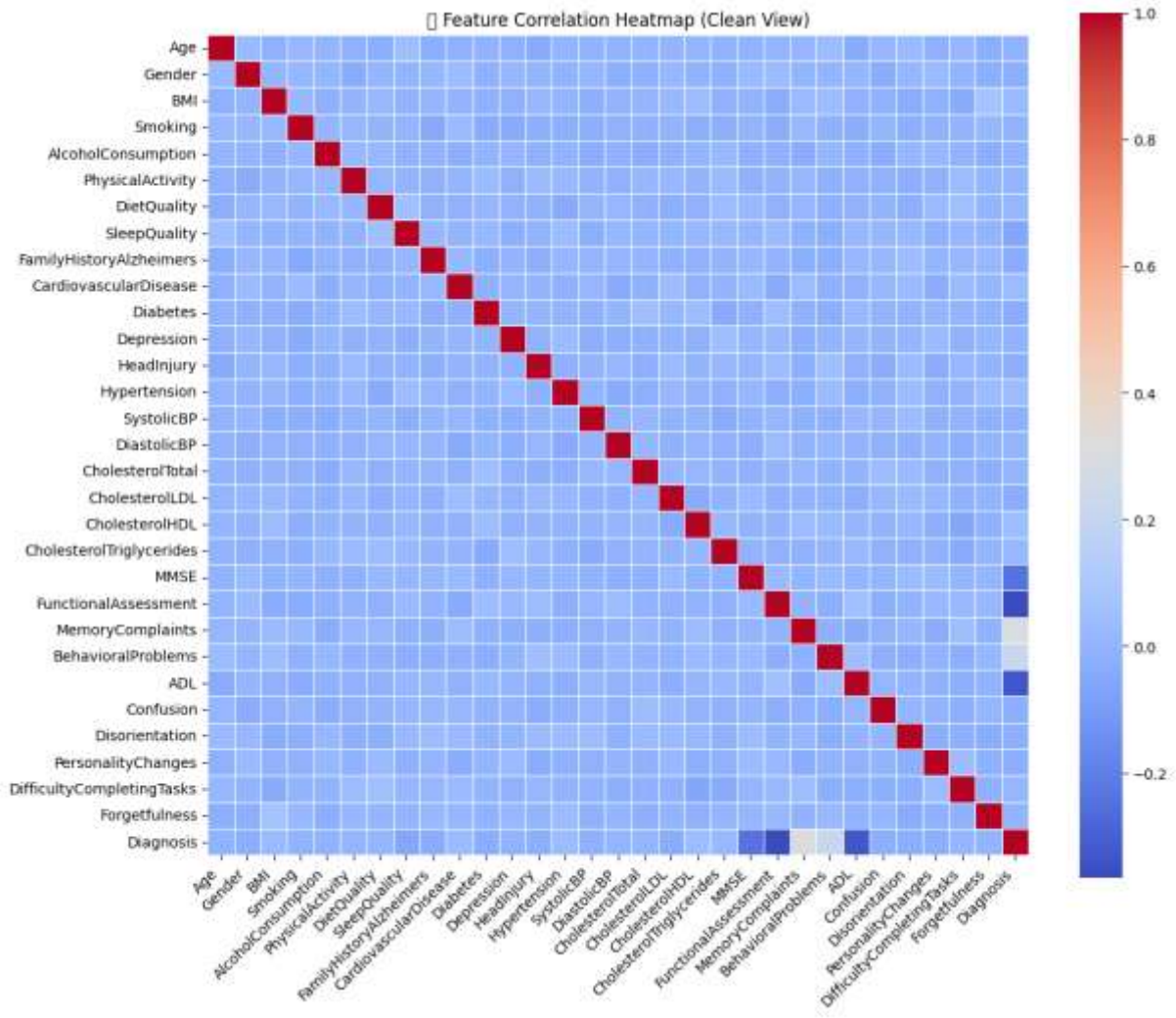
- **Missing Value Handling:** Missing entries in numeric fields (e.g., MMSE scores, ADL scores) were imputed using median values to reduce bias from outliers. Categorical fields were imputed using the mode.
- **Encoding:** Categorical variables such as gender, education level, and diagnosis status were transformed using One-Hot Encoding to enable compatibility with tree-based models and logistic regression.
- **Feature Scaling:** Continuous variables were standardised using StandardScaler to ensure uniform contribution across distance-based models.
- **Outlier Detection:** Z-score analysis was employed to detect and optionally remove extreme values that could distort model training.
- **Target Definition:** The diagnosis variable was mapped to binary labels, where '0' denoted no cognitive impairment and '1' denoted mild cognitive impairment or Alzheimer's diagnosis.

To assess feature interactions and potential multicollinearity, we visualised the Pearson correlation matrix across all selected clinical attributes (Figure 1). This heatmap not only reveals redundant variables but also supports better-informed feature selection for model training.

ISSN: 2583-6129
DOI: 10.55041/ISJEM04812

Figure 1. Feature Correlation Heatmap: Visual representation of Pearson correlation coefficients among selected clinical features. Strong positive correlations are shown in red and strong negative correlations in blue. This aids in identifying redundant variables and potential multicollinearity issues before model training.

Note: ADL – Activities of Daily Living; MMSE – Mini-Mental State Examination; BMI – Body Mass Index.



Visual representation of Pearson correlation coefficients among selected clinical features. Strong positive correlations are shown in red and strong negative correlations in blue. This aids in identifying redundant variables and potential multicollinearity issues before model training.

Note: ADL – Activities of Daily Living; MMSE – Mini-Mental State Examination; BMI – Body Mass Index.

The dataset was split into training (70%) and testing (30%) subsets using stratified sampling to preserve class balance. K-fold cross-validation (k=5) was used within the training set for hyperparameter optimisation.

3.3 Models Implemented

We implemented five machine learning models to evaluate their performance and interpretability in the context of early Alzheimer's prediction. Each algorithm is described below, including its working principle and key mathematical formulation:

3.3.1 Logistic Regression

Logistic regression is a fundamental linear classifier that estimates the probability of the default class (typically 1) using the sigmoid function (Hastie et al., 2001). This function transforms a linear combination of input features into a probability score between 0 and 1, making it suitable for binary classification tasks.

The probability of a positive outcome (y = 1) given input features x is defined as:

P(y=1|x)=11+exp⁡(−(β0+β1x1+β2x2+⋯+βnxn)),$P(y=1|x)=1+\exp(-(\beta 0+\beta 1x1+\beta 2x2+\cdots+\beta nxn))1$,

where β0*β*0 represents the intercept and β1,…,βn*β*1,…,*βn* are the learned coefficients (weights) for each feature x1,…,xn*x*1,…,*xn*, respectively. These coefficients indicate the impact of each feature on the probability of the outcome.

3.3.2 Random Forest

The random forest algorithm is a powerful ensemble learning method that constructs a multitude of decision trees at training time (Breiman, 2001). Each tree in the forest is trained on a different bootstrapped sample of the training data. For classification, the final prediction is determined by a majority vote among the individual trees. This ensemble approach significantly reduces overfitting and improves robustness compared to a single decision tree.

The predicted output y^*y*^ for an input x is given by:

y^=mode(T1(x),T2(x),…,Tk(x)),*y*^=mode(*T*1(x),*T*2(x),…,*Tk*(x)),

where Tk(x)*Tk*(x) denotes the prediction of the k-th decision tree for the input x, and k is the total number of trees in the forest.

3.3.3 XGBoost

XGBoost (eXtreme Gradient Boosting) is an efficient and scalable gradient boosting framework renowned for its high performance (Chen & Guestrin, 2016). It iteratively builds an ensemble of weak learners (typically decision trees) by minimizing a regularized objective function through an additive modeling approach. The regularization term helps control model complexity and prevents overfitting.

The predicted output y^i*y*^*i* for an instance i is the sum of predictions from *K* base functions, fk*fk*:

y^i=∑k=1Kfk(xi),with fk∈F.*y*^*i*=*k*=1∑*Kfk*(x*i*),with *fk*∈F.

The objective function to be minimized is:

Objective=∑i=1nl(yi,y^i)+∑k=1KΩ(fk),Objective=*i*=1∑*nl*(*yi*,*y*^*i*)+*k*=1∑*K*Ω(*fk*),

where l(yi,y^i)*l*(*yi*,*y*^*i*) represents the differentiable loss function measuring the difference between the true label yi*yi* and the predicted label y^i*y*^*i*, and Ω(fk)Ω(*fk*) is the regularization term that penalizes the complexity of the k-th base function fk*fk*.

3.3.4 CatBoost

CatBoost (Categorical Boosting) is an advanced gradient boosting algorithm specifically designed to handle categorical features natively and efficiently (Prokhorenkova et al., 2018). It builds upon the general gradient boosting framework but introduces key innovations to mitigate prediction shift and improve model generalization, particularly with categorical variables (Phani Praveen et al., 2025; Shukla et al., 2023).

Like other gradient boosting methods, CatBoost aims to construct an additive model by iteratively adding new base learners (decision trees) that minimize an objective function. The general form of the prediction y^i*y*^*i* for an instance i after *K* iterations is:

y^i=FK(xi)=FK−1(xi)+ηhK(xi),*y*^*i*=*FK*(x*i*)=*FK*−1(x*i*)+*ηhK*(x*i*),

where FK−1(xi)*FK*−1(x*i*) is the prediction from the ensemble of the previous $K-1$ trees, hK(xi)*hK*(x*i*) is the *K*-th base learner (decision tree) fitted to the negative gradient of the loss function, and ηη is the learning rate.

The objective function minimized is similar to other gradient boosting algorithms:

Objective=∑i=1nL(yi,y^i)+∑k=1KΩ(hk),Objective=*i*=1∑*nL*(*yi*,*y*^*i*)+*k*=1∑*K*Ω(*hk*),

where L(yi,y^i)*L*(*yi*,*y*^*i*) is the loss function (e.g., Logloss for classification, RMSE for regression) measuring the discrepancy between the true value yi*yi* and the current prediction y^i*y*^*i*, and Ω(hk)Ω(*hk*) is a regularization term that penalizes the complexity of the k-th base learner hk*hk*.

3.3.5 TabNet

TabNet is a deep learning architecture uniquely tailored for tabular data (Arik & Pfister, 2021). Unlike traditional deep learning models that often struggle with the heterogeneous nature of tabular inputs, TabNet utilizes sparse attention mechanisms to selectively focus on relevant features at each decision step. This interpretability feature allows for insights into which features are most influential for a given prediction.

The mask M[l]*M*[*l*] at decision step *l* is calculated as:

M[l]=sparsemax(P[l]⊙f[l](x)),*M*[*l*]=sparsemax(*P*[*l*]⊙*f*[*l*](x)),

where P[l]$P[l]$ is the prior scale, f[l](x)$f[l](x)$ is a learnable transformation function applied to the input x, and sparsemax is a sparse version of the softmax function that encourages sparsity in feature selection.

Each model underwent hyperparameter tuning using randomized grid search and 5-fold cross-validation. Model selection was based on optimizing the F1-score and ROC-AUC.

3.4 Explainability

To promote clinical trust and transparency, we applied SHAP (SHapley Additive exPlanations) to all tree-based models (Lundberg & Lee, 2017). SHAP values quantify the contribution of each feature to individual predictions by assigning additive importance scores based on game theory.

- Global Interpretability: SHAP summary plots were generated to visualize the overall impact, frequency, and direction of features across the test set. Features like Functional Assessment, Activities of Daily Living (ADL) status, Mini-Mental State Examination (MMSE) scores, and memory complaints consistently emerged as the most influential.
- Local Interpretability: Force plots and decision plots were used to interpret individual predictions, allowing practitioners to see how different feature values contributed to a specific patient classification.

This interpretability component bridges the gap between model performance and real-world applicability, particularly in high-stakes healthcare settings (Kim et al., 2021).

4.Result

Figures 2 and 3 illustrate the ROC and Precision-Recall curves, respectively. These visualisations affirm the superior threshold performance of ensemble models like CatBoost and XGBoost, with AUCs exceeding 0.94. Notably, CatBoost maintains high precision even at increasing recall values, underscoring its robustness under class imbalance.

Figure 4 displays SHAP summary plots, which interpret model behaviour by highlighting feature contributions. Functional Assessment, MMSE scores, ADL, and Memory Complaints consistently emerged as high-impact variables, confirming their clinical significance and supporting model transparency.

To explore distributional shifts in critical features, Figures 5 and 6 present box plots and violin plots, respectively, showcasing the distribution of key features such as Functional Assessment, ADL, and Memory Complaints across diagnostic categories. The box plots highlight central tendencies and variances between groups, revealing consistently lower median values and wider interquartile ranges among Alzheimer's diagnosed individuals. Violin plots enrich these insights by illustrating multimodal patterns, density variations, and outliers, offering an in-depth look at behavioural markers of cognitive decline.

The confusion matrices in Figures 7a and 7b provide further insights into the classification capabilities of the top-performing models. CatBoost achieved near-perfect sensitivity and specificity, misclassifying only 21 out of 430 test samples. XGBoost displayed similar reliability, reinforcing the consistency of ensemble approaches.

Table 1 summarises model performance. CatBoost achieved the best results (accuracy: 94.9%, F1: 0.95), followed by XGBoost and Random Forest. Logistic Regression underperformed with an F1-score of 0.82. Ensemble models demonstrated strong ROC-AUC scores (0.94+), validated by both ROC and precision-recall curves.

Table 1. Model Performance Metrics

Model	Accuracy	F1 Score	Precision	Recall	ROC-AUC
CatBoost	0.949	0.949	0.960	0.960	0.94
XGBoost	0.948	0.952	0.940	0.950	0.95
Random Forest	0.942	0.941	0.940	0.970	0.94
TabNet	0.881	0.881	0.900	0.910	0.90
Logistic Regression	0.814	0.817	0.910	0.790	0.88

ISSN: 2583-6129
DOI: 10.55041/ISJEM04812

Fgure 2. Receiver Operating Characteristic (ROC) curves comparing the classification performance of all models.

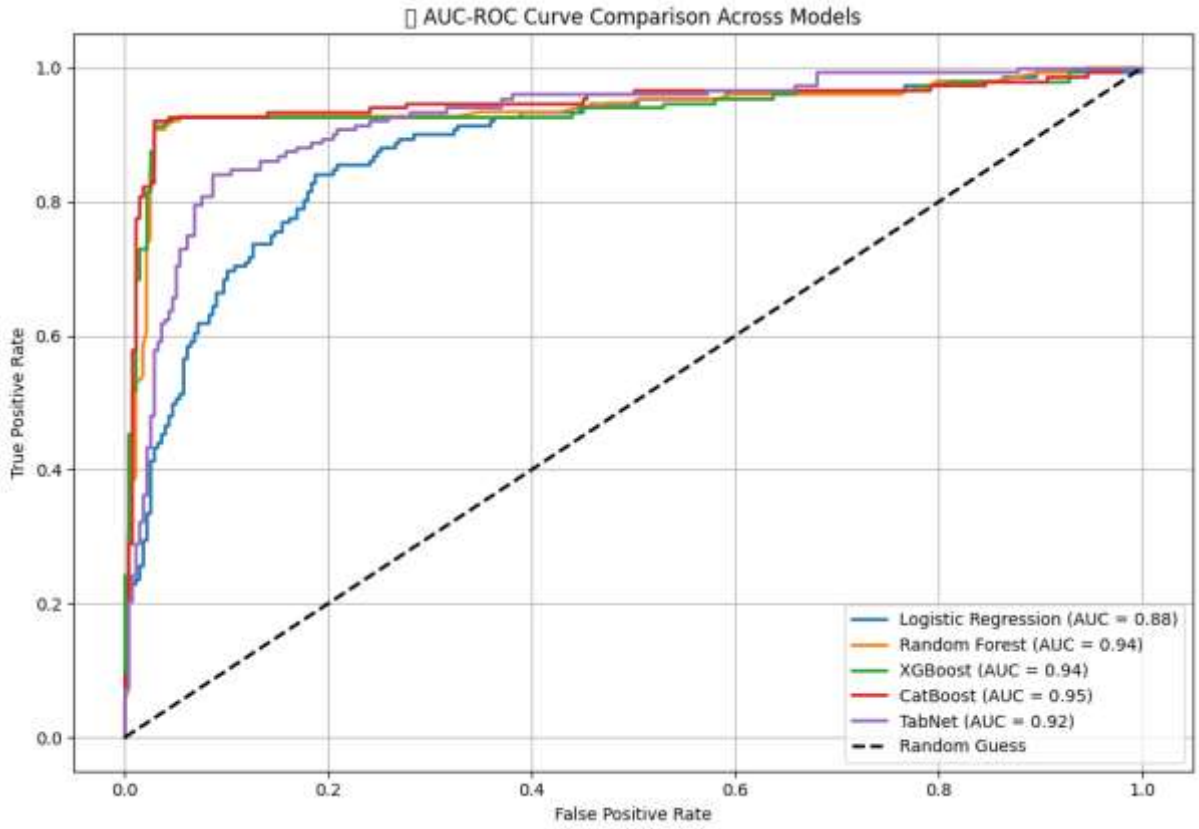


Figure 3. Precision-Recall (PR) curves for all models.

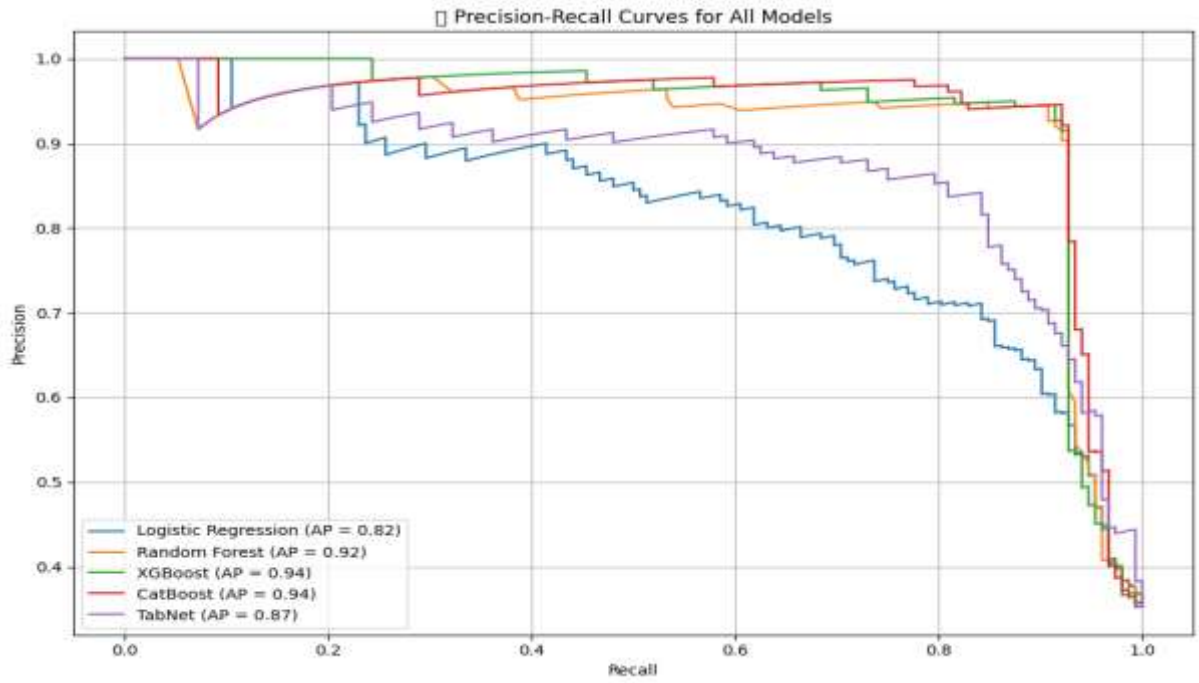


Figure 4. SHAP summary plot for the CatBoost model. Functional Assessment, ADL, and Memory Complaints emerge as dominant predictors.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04812
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

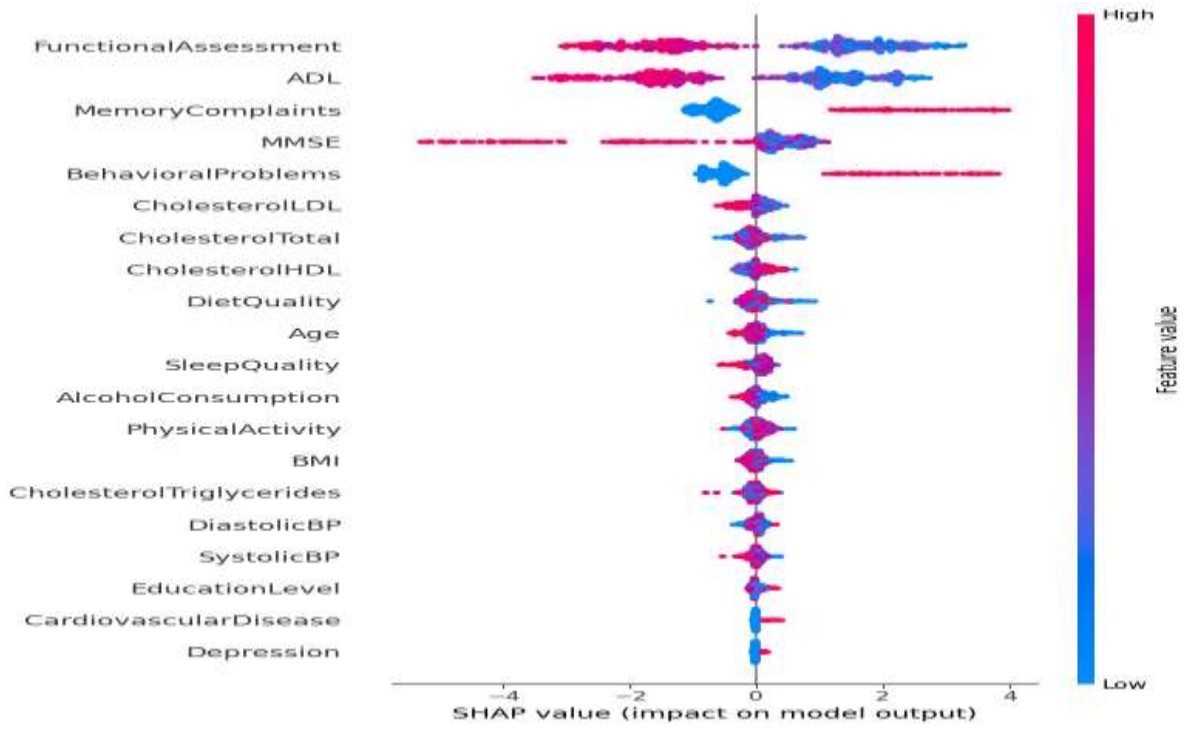


Figure 5. Box plots: Showing the distribution of Functional Assessment, ADL scores, and Memory Complaints across diagnostic groups (0 = Healthy, 1 = Alzheimer's). The visual illustrates feature variance and central tendencies relevant to classification.

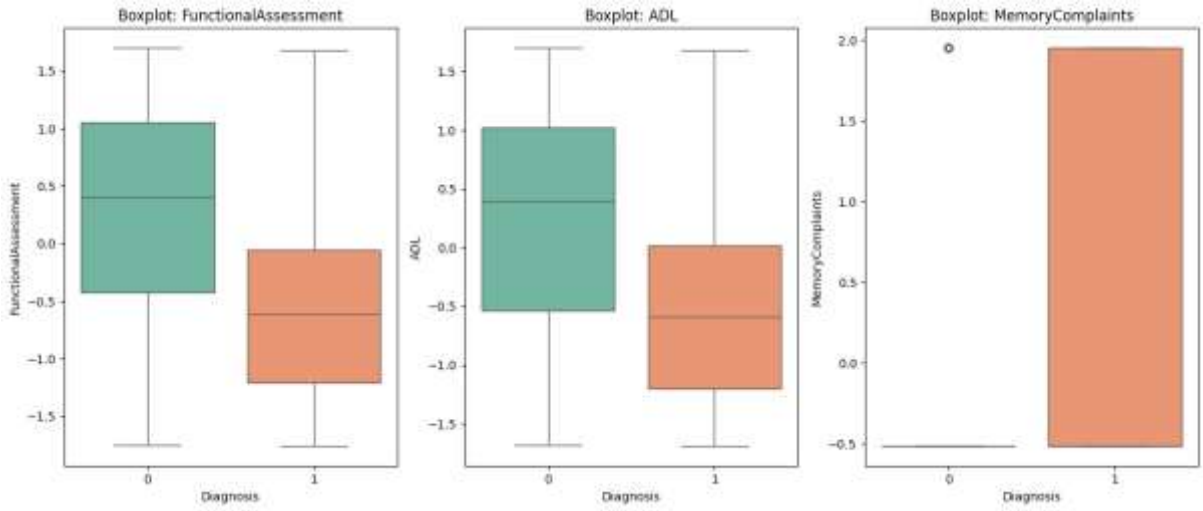


Figure 6. Violin plots: Depicting the distribution density of Functional Assessment, ADL scores, and Memory Complaints by diagnostic category. These plots reveal underlying data spread and outlier behaviour within each class.

International Scientific Journal of Engineering and Management (ISJEM)
Volume: 04 Issue: 07 | July – 2025
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM04812

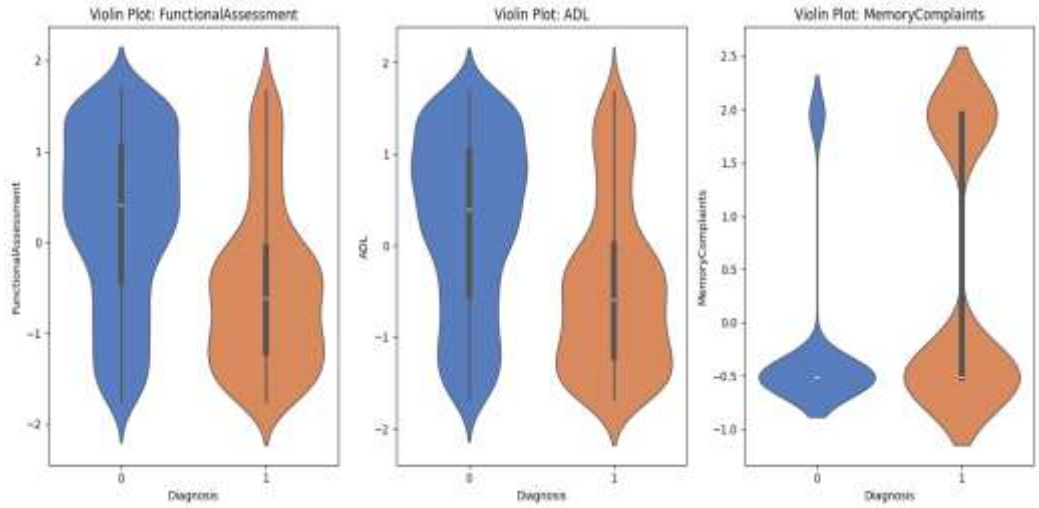


Figure 6a. Confusion Matrix for CatBoost Model

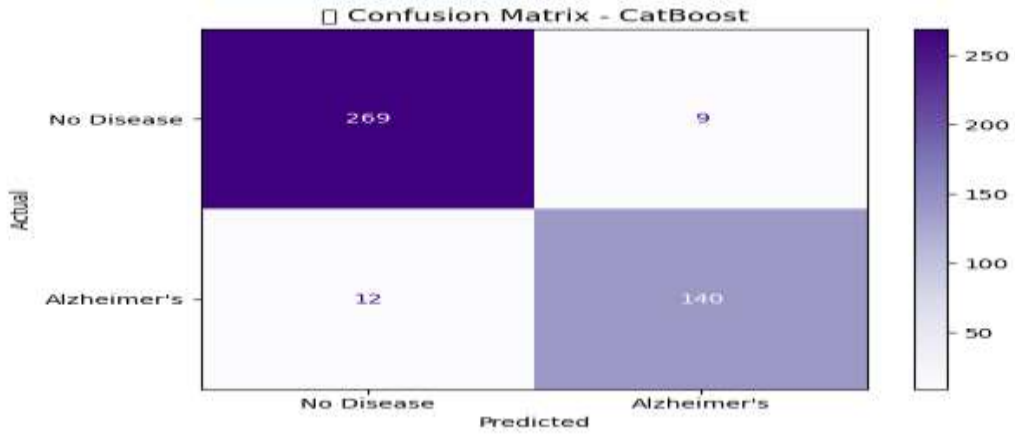
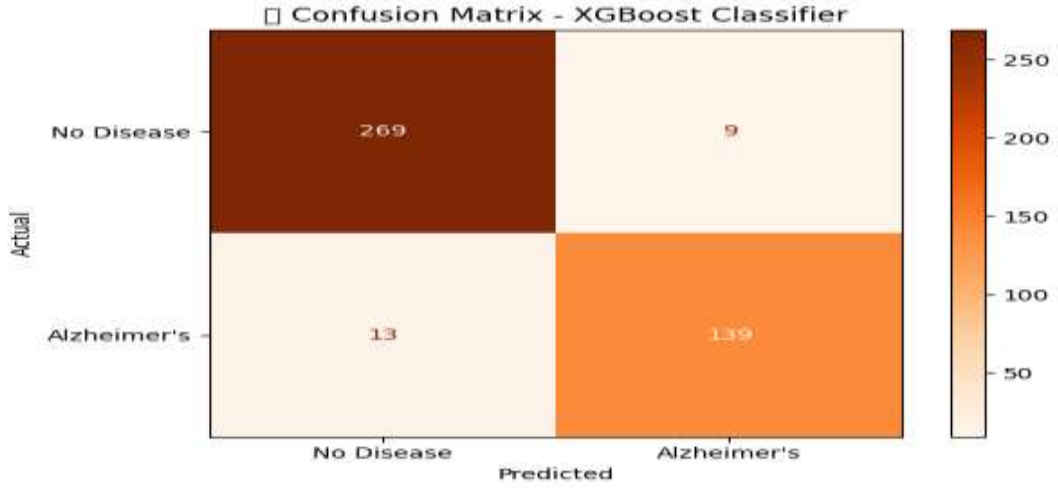


Figure 6b. Confusion Matrix for XGBoost Model



5. Discussion

The performance of CatBoost observed in this study aligns with prior findings where gradient boosting methods have been effective for clinical prediction tasks (Ortiz et al., 2016; Chen & Guestrin, 2016). Its superior performance may be attributed to innovations in

handling categorical data and ordered boosting mechanisms (Prokhorenkova et al., 2018). In comparison, logistic regression showed limitations in modeling nonlinear relationships, a common drawback cited in clinical machine learning applications (Hastie et al., 2001). The TabNet model, while appealing for its attention mechanisms, echoes findings by Arik and Pfister (2021), where performance depended heavily on hyperparameter tuning.

Moreover, the use of SHapley Additive exPlanations (SHAP) values in this study reflects the emphasis placed in recent literature (Linardatos et al., 2020; Lundberg & Lee, 2017) on the need for transparency and interpretability in medical artificial intelligence (AI). Unlike several black-box approaches criticized for their opacity (Hamoud et al., 2025), this framework offers both local and global interpretability, fostering clinician trust. The diagnostic potential demonstrated using only structured clinical inputs is supported by recent studies emphasizing behavioral and cognitive features (Kavitha et al., 2022; Nallapu et al., 2024). In contrast to high-cost neuroimaging studies (Illakiya & Karthik, 2023), this method provides a cost-effective and scalable alternative aligned with healthcare system needs.

Clinical Implications

With its simplicity and explainability, this model could be adapted to real-world screening systems, particularly in resource-limited settings where imaging infrastructure is lacking.

6. Conclusion

In conclusion, this study proposes a clinically meaningful and technically robust machine learning framework for the early detection of Alzheimer's Disease using structured, non-imaging clinical data. Among the models evaluated, CatBoost demonstrated superior predictive accuracy and interpretability, validated through SHAP analysis, which highlighted clinically significant features such as Functional Assessment, ADL scores, and Memory Complaints. The emphasis on explainability ensures transparency, which is vital for clinical adoption of AI-driven tools. Looking forward, expanding the dataset with longitudinal and multi-site records will be critical for improving generalizability. Evaluating algorithmic fairness across demographic factors such as age, gender, and ethnicity is equally important to prevent biased diagnostics. Integration with wearable sensor data could enable continuous, real-time risk assessment, while the deployment of a user-friendly, web-based prediction tool would increase clinical accessibility. Finally, extensive external validation on diverse populations will be essential to translate this framework from research into a dependable decision support system for early Alzheimer's diagnosis.

References

Alzheimer's Disease International. (2023). *World Alzheimer Report 2023: Reducing dementia risk: Never too early, never too late*. https://www.alzint.org/resource/world-alzheimer-report-2023/

Arik, S., & Pfister, T. (2021). Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*. https://ojs.aaai.org/index.php/AAAI/article/view/16826

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Hamoud, A. A., Al-Taweel, D. A., Al-Taweel, F. A., & Al-Taweel, A. A. (2025, March). *Unveiling explainable AI in healthcare: Current trends, challenges, and future directions*. medRxiv. https://www.medrxiv.org/content/10.1101/2024.08.10.24311735v2.full-text

Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer. https://doi.org/10.1007/978-0-387-21606-5

Hosseini-Asl, E., Gimel'farb, G., & El-Baz, A. (2016). *Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network*. arXiv. http://arxiv.org/abs/1607.00556

Illakiya, T., & Karthik, R. (2023). Automatic detection of Alzheimer's disease using deep learning models and neuro-imaging: Current trends and future perspectives. *Neuroinformatics, 21*(2), 339–364. https://doi.org/10.1007/S12021-023-09625-7

Kavitha, C., Mani, V., Srividhya, S. R., Khalaf, O. I., & Tavera Romero, C. A. (2022). Early-stage Alzheimer's disease prediction using machine learning models. *Frontiers in Public Health, 10*, 853294. https://doi.org/10.3389/FPUBH.2022.853294

Kharoua, R. E. (2024). *Alzheimer's disease dataset* [Data set]. Kaggle. https://www.kaggle.com/dsv/8668279. https://doi.org/10.34740/KAGGLE/DSV/8668279

Kim, J. Y., et al. (2021). Development of random forest algorithm based prediction model of Alzheimer's disease using neurodegeneration pattern. *Psychiatry Investigation, 18*(1), 69. https://doi.org/10.30773/PI.2020.0304

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable AI: A review of machine learning interpretability methods. *Entropy, 23*(1), 18. https://doi.org/10.3390/E23010018

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

Nallapu, B. T., et al. (2024, September). *A machine learning approach to predict cognitive decline in Alzheimer's disease clinical trials*. medRxiv. https://doi.org/10.1101/2024.09.03.24312481

Ortiz, A., Munilla, J., Górriz, J. M., & Ramírez, J. (2016). Ensembles of deep learning architectures for the early diagnosis of Alzheimer's disease. *International Journal of Neural Systems, 26*(7). https://doi.org/10.1142/S0129065716500258

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems*. https://proceedings.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html

Ritter, K., Schumacher, J., Weygandt, M., Buchert, R., Allefeld, C., & Haynes, J. D. (2015). Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(2), 206–215. https://doi.org/10.1016/J.DADM.2015.01.006

Sarma, M., & Chatterjee, S. (2025). Machine learning-based Alzheimer's disease stage diagnosis utilizing blood gene expression and clinical data: A comparative investigation. *Diagnostics, 15*(2), 211. https://www.mdpi.com/2075-4418/15/2/211

Shukla, R., Singh, T. R., & Professor, A. (2023). *AlzGenPred: A CatBoost-based method using network features to classify Alzheimer's disease-associated genes from high-throughput sequencing data*. bioRxiv. https://doi.org/10.1101/2023.11.04.565636

World Health Organization. (2025). *Dementia* [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/dementia