Cureus Journal of Computer Science

Open Access Original Article

Implementing a Decision Support Mechanism for Assessing the Probability of Future Alzheimer's Disease Based on the Medical and Demographic Information of Patients

Received 06/17/2025
Review began 06/20/2025
Review ended 10/17/2025
Published 10/22/2025

DOI:
https://doi.org/10.7759/s44389-025-07827-5

Evgenia Psarra [1] , Dimitris Apostolou [1]

1. Department of Informatics, University of Piraeus, Athens, GRC

Corresponding author: Evgenia Psarra, jennypsarra@unipi.gr

Abstract

Alzheimer’s disease (AD) is a progressive neurodegenerative disorder that gradually impairs memory and cognitive function. Early detection of AD is vital, as it allows for timely clinical intervention and potentially slows disease progression. In recent years, machine learning (ML) has emerged as a powerful tool in medical diagnostics, offering support in identifying early indicators of AD. This study investigates the performance of several ML models in classifying early-stage AD using clinical data. Seven ML algorithms were evaluated: Categorical Boosting (CatBoost) and eXtreme Gradient Boosting (XGBoost), which are both types of Gradient Boosting Machine (GBM), Logistic Regression (LR), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Random Forest (RF), and Multi-Layer Perceptron (MLP) Neural Network (NN). These models were compared based on their accuracy, precision, recall, F1-score, specificity, Cohen’s kappa, and Receiver Operating Characteristic (ROC) Curve-Area Under the Curve (AUC), key metrics for assessing classification performance. Among all models tested, CatBoost, which is a GBM-based classifier, achieved the highest overall performance, with an accuracy of 96.74%, precision of 96.86%, recall of 94.48%, F1-score of 95.65%, specificity of 98.13%, Cohen’s kappa of 93.05%, and ROC-AUC of 96.30%. RF also performed strongly, with 95.81% accuracy and 94.34% F1-score. XGBoost, another GBM classifier, showed comparable effectiveness with 95.58% accuracy and 94.08% F1-score. In contrast, traditional classifiers like LR and SVM delivered lower accuracy (84.42% and 83.72%, respectively) and F1-scores near 77%. MLP outperformed these traditional methods with 85.58% accuracy and 78.91% F1-score. KNN had the weakest performance, particularly in recall (49.08%) and F1-score (62.26%), despite relatively high precision (85.11%) and specificity (94.76%). Overall, the findings affirm that the models CatBoost and RF provide the most reliable and generalizable solutions for the classification task in this study, exhibiting high discriminative capability and stability across all evaluated metrics.

Categories: AI applications, Health Informatics, Machine Learning (ML)
Keywords: alzheimer disease, medical decision support system, artificial neural networks, machine learning, gradient boosting machine, logistic regression, support vector machine, k-nearest neighbors, random forest, multi-layer perceptron

Introduction

Alzheimer’s disease (AD) is a progressive neurodegenerative disorder marked by irreversible cognitive impairment [1]. Based on the work of Rony et al. [2], one growing concern, made even more pressing by the COVID-19 pandemic, is the increased susceptibility of older adults to mental health issues. Conditions such as depression and anxiety are all too often dismissed as just a natural part of getting older, when in fact they are both common and serious. Adding to this are neurodegenerative illnesses, such as the AD, which bring about memory loss, confusion, and profound changes in personality. Despite the significant impact these conditions can have, they frequently go undetected or remain untreated in older populations, underscoring the complexity and urgency of addressing mental health in later life. AD can be prevented by consuming vitamins such as Omega-3 [3] or vitamin D [4].

Early detection remains a significant challenge, as clinical symptoms often emerge long after the initial onset of pathology. This underscores the importance of identifying subtle alterations in biomarkers, which are frequently observable through neuroimaging techniques. In this context, computer-aided diagnostic systems employing machine learning (ML) and deep learning have emerged as promising tools for the early identification of diseases such as AD [1], type 2 diabetes [5], liver disease [6], or heart stroke [7], by analyzing complex and heterogeneous data inputs [1]. Kaur and Sachdeva [1] in their review of recent studies highlight the adoption of diverse algorithmic frameworks, with a strong emphasis on practical aspects such as data acquisition methods, input modalities, feature engineering, and model validation. Their comparative performance analysis suggests that deep learning models, especially those designed to integrate multimodal information, demonstrate notable efficacy in early-stage detection and multiclass classification of AD.

How to cite this article
Psarra E, Apostolou D (October 22, 2025) Implementing a Decision Support Mechanism for Assessing the Probability of Future Alzheimer's Disease Based on the Medical and Demographic Information of Patients. Cureus J Comput Sci 2 : es44389-025-07827-5. DOI https://doi.org/10.7759/s44389-025-07827-5

According to Priya et al. [8], AD is a progressive neurodegenerative disorder that impairs cognitive functions and gradually deteriorates neuronal structure and performance. Early detection of AD allows for prompt intervention, which can enhance patient outcomes and overall quality of life. In this context, ML plays a crucial role in the diagnostic process, assisting healthcare providers in identifying abnormalities and prioritizing critical cases. As a result, significant research has recently focused on developing effective methods for the prediction and early detection of AD.

In the study by Mao et al. [9], weighted gene co-expression network analysis combined with various ML algorithms was utilized to pinpoint critical palmitoylation-associated genes linked to AD. Further analyses, including immune cell infiltration profiling and construction of a competing endogenous RNA (ceRNA) network, highlighted the potential involvement of ZDHHC22 gene in immune modulation and gene regulatory processes, offering novel perspectives on the underlying mechanisms of AD pathogenesis.

According to Jiao et al. [10], AD is characterized by intricate disruptions in various biological pathways, underscoring the importance of comprehensive blood-based biomarkers for accurate and early diagnosis. They developed a cost-effective and accessible ML approach utilizing digital biomarkers derived from plasma spectral data to distinguish individuals with AD or mild cognitive impairment from healthy controls, as well as to differentiate AD from other neurodegenerative disorders. Retrospective data from 1,324 participants, comprising 293 amyloid beta-positive AD cases, 151 with mild cognitive impairment, 106 with dementia with Lewy body dementia, 106 with frontotemporal dementia, 135 with progressive supranuclear palsy, and 533 healthy controls, were collected between July 2017 and August 2023. Feature selection techniques and a Random Forest (RF) classifier were employed to identify relevant digital biomarkers. The extracted features from Attenuated Total Reflectance-Fourier Transform Infrared plasma spectra effectively captured AD-related pathological alterations, supporting their potential as digital biomarkers for early detection and differential diagnosis of AD.

According to Javed et al. [11], chronic diseases pose a significant global health burden. In recent years, deep ML algorithms have been increasingly applied to the diagnosis of such conditions. Early diagnosis and intervention are critical, as they reduce the risk of disease progression and associated mortality. Their study presents a deep learning-based approach that achieves improved diagnostic accuracy, with their model reaching 96.06% accuracy. These findings contribute to the advancement of personalized Healthcare 5.0 by enabling more efficient prediction of chronic diseases. Furthermore, their results underscore the importance of selecting appropriate models tailored to specific disease types, particularly when leveraging data from internet of medical things-enabled devices.

This work builds upon our previous work [12-21], where we integrated ML mechanism within an access control framework by using patient's contextual information, enhancing the overall security and functionality of electronic health information systems. Our work aims: i) to study various ML techniques applied to AD prediction and based on a medical dataset consisting of medical records of 2,149 patients, and ii) to compare the findings of our study with other ML techniques in the literature.

In this work, six ML algorithms were evaluated for their effectiveness in detection of AD: Gradient Boosting Machine (GBM), Logistic Regression (LR), Support Vector Machines (SVMs), k-Nearest Neighbors (KNN), RF, and Multi-Layer Perceptron (MLP) Neural Network (NN) method by using the medical dataset "Alzheimer's Disease Dataset" [22] derived from 2,149 patients. The rest of this work is organized as follows: i) The "Introduction" reviews previous research that has applied ML techniques to AD detection, using either the same dataset or similar ones. ii) The "Materials and Methods" section describes the dataset and the ML methods used in this study. iii) In the "Results", we present the performance metrics of each model, along with a comparison to existing studies that used the same dataset with different approaches. iv) The "Discussion" explores how our findings align with or differ from past research, and also compares the strengths and limitations of the various ML methods. v) Finally, the "Conclusion" summarizes the main findings and highlights the contributions of this work.

Suñé [23] reported accuracies of ML models for AD prediction, by implementing the "Alzheimer's Disease Dataset" [22], which range from 63% using SVMs to 91.8% with convolutional NNs. The SVM models achieved accuracies between 80% and 88%. His enhanced performance appears to stem from the integration of both medical and lifestyle variables, indicating that multimodal data can significantly improve diagnostic accuracy with SVMs. In contrast, the NN models, which did not incorporate convolutional layers, attained accuracies between 63% and 86%, falling short of the best results in the literature. Despite employing deep learning architectures with varying hidden layers, the absence of convolutional components may have limited their ability to extract more complex patterns necessary for top-tier performance. Among all models tested, the Decision Tree (DT) classifier emerged as the top performer, achieving an accuracy of 93.4% when trained on statistically significant variants and normalized using MinMax scaling, surpassing previously reported state-of-the-art results. Closely following was the RF classifier, with an accuracy of 93.37%, under the same configuration. Additional strong performances were observed from both DT and RF models using MinMax scaling, each achieving 92% accuracy. These findings suggest that MinMax normalization pairs particularly well with tree-based algorithms, as all top-four performing models employed this combination. Further supporting this trend,

two additional tree-based models using Standard Scaler also surpassed the 90% accuracy threshold. Tree algorithms are particularly well-suited for this domain, as they inherently filter out non-informative features, a critical advantage when working with heterogeneous medical data. In the 80% accuracy range, notable models include SVMs with significant variant selection and MinMax scaling, NNs with significant features and Standard scaling, and KNN using the same approach. LR models also performed consistently in this range, despite their simplicity.

The results of the study by Balbal [24] underscore the strong potential of ML algorithms in accurately predicting AD. His analysis was conducted using a publicly available dataset from the Kaggle platform, which is used for our work as well [22] comprising 2,149 patient records and 35 demographic and clinical features, all of which were complete and free from missing values. This ensured the dataset's integrity and suitability for robust ML analysis. The primary objective of the study was to evaluate the predictive performance of several widely used ML algorithms that are commonly applied in disease prediction and classification research.

The algorithms assessed included eXtreme Gradient Boosting (XGBoost), KNN, Gaussian Naive Bayes, DT, RF, AdaBoost, and SVMs. Their performance was evaluated using four key metrics: accuracy, precision, recall, and F1-score. Among these, XGBoost outperformed all other models, achieving the highest accuracy of 95.35%, reflecting its strong capacity to model complex relationships within multidimensional clinical data. This aligns with existing literature, where XGBoost classifiers have consistently demonstrated high effectiveness in similar predictive tasks.

Conversely, KNN yielded the lowest accuracy at 75.54%, although it still performed above a baseline threshold, indicating a general adequacy of all models in predicting AD to varying degrees. Despite differences in performance, the study's findings confirm the overall efficacy of ML algorithms in supporting early diagnosis efforts for AD. In particular, the outstanding performance of XGBoost highlights the value of advanced ensemble techniques in leveraging intricate data patterns for improved clinical decision-making.

A study by Sendhil et al. [25] presents a comprehensive overview of traditional approaches employed in the detection and treatment of AD, by using the "Alzheimer's Disease Dataset" that we used in our work as well [22]. Utilizing a dataset comprising 2,149 individuals with varied demographic, lifestyle, clinical, and cognitive characteristics, five ML algorithms were evaluated and compared. Among these, KNN and the RF classifier demonstrated the highest performance, each achieving accuracy rates exceeding 92% along with well-balanced F1-scores. Their results underscore the robustness and reliability in the context of early-stage AD prediction, outperforming the other models assessed in their study.

A study by Airlangga [26] evaluates the predictive performance of three ML architectures, namely, MLP NN, convolutional NN, and long short-term memory, for the detection of AD using the same multidimensional dataset, which was used in our work as well. Among the models, convolutional NN outperformed the others, achieving an average accuracy of 88.65%, compared to 84.41% for MLP and 75.57% for long short-term memory. These findings emphasize convolutional NN's strength in capturing spatial dependencies within clinical data, positioning it as a highly effective approach for AD diagnosis. In contrast, the long short-term memory model underperformed, likely due to the dataset's lack of temporal continuity, which limited its ability to leverage sequence-based learning. Overall, the study highlights the critical importance of aligning model architecture with the nature of the dataset and lays the groundwork for incorporating ML techniques into clinical decision-making processes.

According to Raghupathy et al. [27], ensemble classifiers integrated with explainable artificial intelligence have become a valuable tool in diagnosing AD. Specifically, boosting-based ensemble techniques paired with shapley additive explanations provide a robust framework for AD diagnosis. This study explores the use of boosting ensemble ML models like XGBoost, LightGBM, and Gradient Boosting (GB) in conjunction with shapley additive explanations for feature selection. The proposed approach demonstrated high effectiveness, achieving over 94% accuracy using a minimal set of features in the detection process.

The framework of Oncu [28] integrates two distinct artificial intelligence models. The artificial NN was trained using clinical data from 1,200 patients, encompassing 31 features related to demographics, symptoms, and behavior, to evaluate the risk of AD. The dataset that the authors used for their artificial NN model was used in our work as well. In parallel, a convolutional NN processed 4,876 MRI scans to confirm diagnoses and categorize the disease into four stages: non-demented, very mild, mild, and moderate dementia. The dataset that they used for the convolutional NN model was different from the one that we used in our work. To improve transparency and clinical relevance, Gradient-weighted Class Activation Mapping visualizations were used to highlight significant imaging features. The artificial NN offers a fast and accessible tool for estimating AD risk based on clinical inputs, while the convolutional NN verifies the diagnosis and assesses the severity through imaging analysis. Combined the methods above, they form a robust diagnostic approach that merges clinical and imaging data for enhanced precision. Their artificial NN reached an accuracy of 87.08% (using the "Alzheimer's Disease Dataset," which is the

same dataset that we used in our work) in detecting high-risk individuals. Additionally, their convolutional NN achieved 97% accuracy in staging the disease (using the "Augmented Alzheimer MRI Dataset," which is different from our work).

Materials And Methods

Six ML algorithms were evaluated for their effectiveness in the early detection of AD: RF, KNN, SVMs, LR, GBM, and MLP NN method. These models were compared based on their accuracy, precision, recall, F1-score, specificity, Cohen's kappa, and receiver operating characteristic (ROC) curve-area under the curve (AUC), key metrics for assessing classification performance.

Random forest classifier

The RF algorithm is a widely used ensemble learning technique that aggregates the outputs of multiple DTs to derive a single predictive outcome. One of its strengths lies in the ability to assess feature importance at both the model level and on a per-instance basis. Since it evaluates various combinations of features across its trees, the contribution of each feature to prediction accuracy can be estimated by measuring the increase in error when that feature is excluded or permuted.

K-nearest neighbors

KNN classifies new instances by assigning them to the most common class among their nearest neighbors in the feature space. In the context of AD detection, this straightforward method has been utilized to identify early stages of cognitive decline, ranging from no impairment to very mild, mild, and moderate cognitive loss, based on similarity to previously classified patient data.

Support vector machines

SVM is a supervised learning method used for both classification and regression tasks. It works by identifying the optimal hyperplane that separates data points from different classes in a high-dimensional space. This method is particularly effective in handling complex and high-dimensional datasets.

Logistic regression

LR is a statistical model designed for binary classification, estimating the probability that a given input belongs to one of two possible categories. In the context of AD prediction, it has been employed to build risk models based on variables such as cognitive performance, demographic factors, and health indicators. This model estimates the likelihood of developing AD by analyzing the relationship between these predictors and the log-odds of disease presence.

Gradient boosting machine (CatBoost and XGBoost)

GBM is an ensemble method that enhances prediction performance by iteratively combining weak learners, typically decision trees. Each iteration focuses on correcting the errors made by the previous model, allowing the algorithm to gradually refine its predictions. GBM is particularly effective at identifying subtle patterns in the data that may be indicative of early AD progression. More precisely, two types of GBM are utilized in this work: CatBoost and XGBoost. CatBoost is a GBM algorithm designed to handle categorical data efficiently and it aims to make gradient boosting less prone to overfitting. XGBoost is a highly efficient gradient boosting library and it is suitable for large scale datasets.

Multi-layer perceptron neural network

Artificial NNs are composed of three main types of layers: input, hidden, and output layers. The number of hidden layers determines the depth of the network. In fully connected NNs, each neuron in one layer is connected to every neuron in the following layer. The output from one layer serves as the input to the next, processed through an activation function applied to each neuron. Moreover, each connection is assigned a specific weight, and a bias term is associated with each layer. The output of each layer can be represented by the following equations:

$$a_l = (W_l h_{l-1}) + b_l$$

$$h_l = f(a_l)$$

In this context, l denotes the index of the layer, f refers to the activation function (such as softmax, ReLU, or sigmoid), W_l represents the weight matrix of layer l, h_l is the activation output, and b_l signifies the bias term for that specific layer. The equation below presents the loss function, where N is the number of nodes in the network's output layer, $P(a)$ is the predicted output, α stands for the NN's parameters, and y corresponds to the true target value.

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - P(a)^2)$$

The primary goal in training a NN is to optimize its parameters by minimizing the loss function. The equation below presents the loss function with an added regularization term, denoted as $\Omega(a)$, which serves as a penalty factor to prevent overfitting.

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - P(a))^2 + \Omega(a)$$

Performance indices

In this study, we initially assessed the performance of the developed models using accuracy as the primary evaluation metric, as defined in the following equation:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

True Positives (TP) and True Negatives (TN) correspond to correctly classified instances, while False Positives (FP) and False Negatives (FN) represent misclassified outcomes. In addition to accuracy, the following performance evaluation metrics were also employed to comprehensively assess the model's effectiveness:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision}+\text{Recall}} = \frac{2*TP}{2*TP+FP+FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Cohen's kappa} = \frac{2 * (TP * TN - FN * FP)}{(TP + FP) * (FP + TN) + (TP + FN) * (FN + TN)}$$

Role of ML models

According to Sendhil et al. [25], the application of these ML models provides an effective toolkit for enhancing the accuracy and efficiency of AD detection. They not only streamline the diagnostic process but also accelerate decision-making. For instance, GBM's iterative nature enables it to uncover nuanced patterns associated with disease onset. LR offers a rapid and interpretable means of assessing risk based on structured health data. SVM excels at classifying different risk levels through its robust decision boundaries, while KNN provides a simple yet effective means of stratifying individuals based on cognitive decline. Finally, RF, which consists of a set of DTs, improves diagnostic reliability by leveraging ensemble learning to reduce overfitting and increase generalization. Collectively, these models contribute significantly to the development of automated, data-driven systems capable of supporting clinicians in the early identification of AD, ultimately aiding in timely intervention and improved patient outcomes.

The implementation of our models is divided into the following stages: 1) Data Preprocessing (which comprises the following sub-steps: a) Checking data integrity, b) Dropping the unnecessary columns "PatientID" and "DoctorInCharge" from the dataset, c) Normalizing the columns with MinMaxScaler, d) Standardizing the columns with StandardScaler), 2) Splitting data into features (32 fields) and target (1 field of "Diagnosis"), 3) Splitting data, after shuffling, into training and testing sets (80% for training set and 20% for testing one), 4) Leveraging the appropriate Classifier (which comprises the following respective methods: a) Random ForestClassifier, b) k-Nearest Neighbors classifier, c) Support Vector classifier, d) Logistic Regression classifier, e) XGBoost classifier, f) CatBoost classifier, g) MLP Neural Network classifier), 5) Training the diagnosis model, 6) Saving the diagnosis model, 7) Inferencing by the trained model (from X_test set predicting the equivelent Y_pred set), 8) Calculating the Confusion Matrix from equivalent Y_test and Y_pred sets (deriving the terms TP, TN, FP and FN), 9) Composing the Performance Indices through Confusion Matrix terms, 10) Hyperparameters tuning for optimizing the performance of each classifier.

For all classifiers, the sklearn module was used, except for the MLP NN classifier where the Keras API of the TensorFlow platform was adapted.

The architecture of the MLP NN is illustrated in Figure *1*.

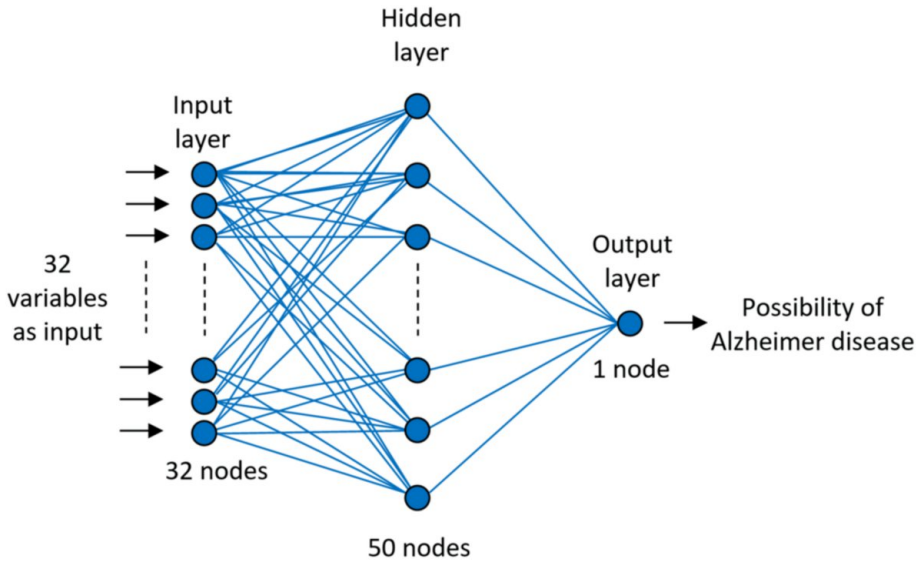

FIGURE 1: Architecture of multi-layer perceptron neural network

More specifically this neural network consists of the following layers: i) a 32-node input layer with Relu activation function, ii) a 50-node hidden layer with Relu activation function, and iii) an output layer with a single node and a Sigmoid activation function. The neural network uses the Adam optimizer and the 'binary_crossentropy' loss function. This model was trained for 25 epochs and a batch size of 32.

The dataset of our study

The dataset of our study [22] contains medical, demographic, lifestyle, functional, and cognitive information for 2,149 patients who may suffer from AD. An indicative part of this dataset is illustrated in Figure *2*.

Patient ID	Age	Gender	Ethnicity	Education Level	BMI	Smoking	Alcohol Consumption	Physical Activity	Diet Quality	Sleep Quality	Family History Alzheimers	Cardiovascular Disease	Diabetes	Depression	Head Injury	Hypertension	Systolic BP	Diastolic BP
4751	73	0	0	2	22.93	0	13.29721773	6.32711	1.3472	9.0257	0	0	1	1	0	0	142	72
4752	89	0	0	0	26.83	0	4.542523818	7.61988	0.5188	7.1513	0	0	0	0	0	0	115	64
4753	73	0	3	1	17.8	0	19.55508453	7.84499	1.8263	9.6736	1	0	0	0	0	0	99	116
4754	74	1	0	1	33.8	1	12.20926555	8.428	7.4356	8.3926	0	0	0	0	0	0	118	115
4755	89	0	0	0	20.72	0	18.45435609	6.31046	0.7955	5.5972	0	0	0	0	0	0	94	117
4756	86	1	1	1	30.63	0	4.140143784	0.21106	1.5849	7.262	0	0	1	0	0	0	168	62
4757	68	0	3	2	38.39	1	0.646047271	9.25769	5.8974	5.4777	0	0	0	0	1	0	143	88
4758	75	0	0	1	18.78	0	13.72382571	4.64945	8.3419	4.2132	0	0	0	0	0	0	117	63
4759	72	1	1	0	27.83	0	12.16784763	1.53136	6.7369	5.7482	0	0	0	0	0	1	117	119
4760	87	0	0	0	35.46	1	16.02868824	6.44077	8.086	7.5518	0	1	0	0	0	0	130	78
4761	89	0	3	1	39.46	0	9.811292129	8.81995	0.434	7.6441	0	0	0	0	0	0	131	86
4762	78	0	0	2	22.46	1	19.30018298	3.83464	8.2792	8.3123	0	0	1	0	0	1	165	97
4763	84	1	0	1	26.77	0	10.97802164	3.97808	7.0244	8.253	1	0	0	1	0	0	145	63
4764	78	1	0	1	28.87	1	10.1947063	0.63128	1.6533	7.3332	1	0	0	0	0	0	137	82

Cholesterol Total	Cholesterol LDL	Cholesterol HDL	Cholesterol Triglycerides	MMSE	Functional Assessment	Memory Complaints	Behavioral Problems	ADL	Confusion	Disorientation	Personality Changes	Difficulty Completing Tasks	Forgetfulness	Diagnosis
242.36684	56.150897	33.682563	162.189143	21.46353	6.51887697	0	0	1.725883	0	0	0	1	0	0
231.162595	193.407996	79.028477	294.630909	20.61327	7.1186955	0	0	2.592424	0	0	0	0	1	0
284.181858	153.322762	69.772292	83.6383241	7.356249	5.89507735	0	0	7.119548	0	1	0	1	0	0
159.58224	65.3666368	68.457491	277.577358	13.99113	8.9651063	0	1	6.481226	0	0	0	0	0	0
237.602184	92.8696999	56.874305	291.19878	13.51761	6.04503877	0	0	0.014691	0	0	1	1	0	0
280.712539	198.334629	79.080503	263.943655	27.51753	5.51014408	0	0	9.015686	1	0	0	0	0	0
263.734149	52.4706696	66.533369	216.489175	1.964413	6.06212447	0	0	9.236328	0	0	0	0	1	0
151.383137	69.6235104	77.346816	210.570866	10.13957	3.40137351	0	0	4.517248	1	0	0	0	1	1
233.605755	144.04574	43.075893	151.164186	25.82073	7.39606098	0	1	0.756232	0	0	1	0	0	0
281.63005	130.49758	74.291247	144.175975	28.38841	1.14890353	0	1	4.554394	0	0	0	0	0	0
224.526437	160.756475	68.328313	172.78522	19.69611	8.24696812	0	0	2.01894	0	1	0	0	1	0
254.58656	132.960012	39.009282	344.448595	21.20562	5.56854144	0	0	5.467267	0	1	0	1	1	0
254.435386	148.291423	98.585547	81.525137	2.20034	0.51513642	0	0	9.003759	1	0	0	0	0	0
221.305338	194.600379	26.33392	357.582771	7.852082	4.51071286	1	0	1.939596	0	1	0	0	0	1

FIGURE 2: Indicative health records of our dataset

ADL, Activities of Daily Living; BMI, Body Mass Index; BP, Blood Pressure; HDL, High-Density Lipoprotein; LDL, Low-Density Lipoprotein; MMSE, Mini-Mental State Examination

More specifically, this dataset includes the following fields:

1) Patient ID (identifier of each patient), 2) Age (ranges from 60 to 90), 3) Gender (0 for a male patient and 1 for a female), 4) Ethnicity (with indices 0: Caucasian, 1: African American, 2: Asian, 3: Other), 5) Education Level (0: None, 1: High School, 2: Bachelor's, 3: Higher), 6) BMI (Body Mass Index, from 15 to 40), 7) Smoking (0: no smoking, 1: yes), 8) Alcohol Consumption (Weekly consumption in units, from 0 to 20), 9) Physical Activity (Weekly physical activity in hours, from 0 to 10), 10) Diet Quality (Diet quality score, from 0 to 10), 11) Sleep Quality (Sleep quality score, ranging from 4 to 10), 12) Family History Alzheimer's (Family history of AD, 0: No , 1: Yes), 13) Cardiovascular Disease (Presence of cardiovascular disease, 0: No , 1: Yes), 14) Diabetes (Presence of diabetes, 0: No , 1: Yes), 15) Depression (Presence of depression, 0: No , 1: Yes), 16) Head Injury (History of head injury, 0: No , 1: Yes), 17) Hypertension (Presence of hypertension, 0: No , 1: Yes), 18) SBP (Systolic Blood Pressure, 90 to 180 mmHg), 19) DBP (Diastolic Blood Pressure, 60 to 120 mmHg), 20) Cholesterol Total (Total cholesterol levels, 150 to 300 mg/dL), 21) Cholesterol LDL (Low-density lipoprotein cholesterol levels, 50 to 200 mg/dL), 22) Cholesterol HDL (High-density lipoprotein cholesterol levels, 20 to 100 mg/dL), 23) Cholesterol Triglycerides (Triglycerides levels, 50 to 400 mg/dL), 24) MMSE (Mini-Mental State Examination score, from 0 to 30), 25) Functional Assessment (Functional assessment score, 0 to 10), 26) Memory Complaints (Presence of memory complaints, 0: No , 1: Yes), 27) Behavioral Problems (Presence of behavioral problems, 0: No , 1: Yes), 28) ADL (Activities of Daily Living score, 0 to 10), 29) Confusion (Presence of confusion, 0: No , 1: Yes), 30) Disorientation (Presence of disorientation, 0: No , 1: Yes), 31) Personality Changes (Presence of personality changes, 0: No , 1: Yes), 32) Difficulty Completing Tasks (Presence of difficulty completing tasks, 0: No , 1: Yes), 33) Forgetfulness (Presence of forgetfulness, 0: No , 1: Yes), 34) Diagnosis (Diagnosis status for AD, 0: No , 1: Yes), and 35) Doctor In Charge (code-name of doctor in charge, "XXXConfid" for all).

As illustrated in Figure *2*, each patient is identified by a numeric identifier, for data confidentiality purposes. In addition, the 35th field of this medical dataset titled “Physician in Charge” displays the code name “XXXConfid” for all patients, for the same privacy reasons.

The distribution of healthy persons along with the ones suffering from AD is shown in Table *1*.

	Healthy	Alzheimer's Disease	Total
Number of Patients	1,389	760	2,149
%	64.7%	35.3%	100%

TABLE 1: Comparison of population characteristics: Healthy controls vs. Alzheimer's disease subjects

As seen in Table *1*, the dataset used in this study comprises a total of 2,149 samples, with 1,389 labeled as healthy (64.7%) and 760 labeled as AD cases (35.3%). It should be noted that before utilizing the dataset, we conducted data normalization and data standardization.

Additionally, the distribution of the dataset fields about Ethnicity, Gender, Education Level, Age, BMI, and SBP is demonstrated in Tables *2*-*7*.

Ethnicity	Caucasian	African American	Asian	Other
Number of Patients	1,278	454	206	211

TABLE 2: Distribution of 2,049 patients according to Ethnicity

Gender	Male	Female
Number of Patients	1,061	1,088

TABLE 3: Distribution of 2,049 patients according to Gender

Education Level	None	High School	Bachelor	Higher
Number of Patients	446	854	636	213

TABLE 4: Distribution of 2,049 patients according to Education Level

Age	60-69	70-79	80-90
Number of Patients	701	703	745

TABLE 5: Distribution of 2,049 patients according to Age

BMI	Underweight (BMI < 18.5)	Healthy Weight (18.5 ≤ BMI ≤ 24.9)	Overweight (25.0 ≤ BMI ≤ 29.9)	Obese (BMI ≥ 30)
Number of Patients	297	534	433	875

TABLE 6: Distribution of 2,049 patients according to Body Mass Index (BMI)

SBP (mm Hg):	90-119	120-139	140-180
Number of Patients	729	476	944

TABLE 7: Distribution of 2,049 patients according to average Systolic Blood Pressure (SBP)

In our study, in order to assess the probability of a patient who may suffer from AD, we implemented the following methods: 1) RF by leveraging the RF classifier, 2) KNN by leveraging the KNN classifier, 3) SVM by leveraging the Support Vector Classifier (SVC), 4) LR by leveraging the LR classifier, 5) GBMs, by leveraging the XGBoost and the CatBoost classifiers, and 6) MLP NNs by leveraging the MLP classifier.

Results And Discussion

Parameterization

Initially, we thoroughly studied each classification method by simultaneously tuning some main parameters of each classifier so as to optimize their performance. All experiments were based on the performance index of accuracy, and they all are shown in a compact Table *8*.

Random Forest classifier							
n_estimators parameter	**10**	**50**	**100**	**200**	**300**	**310**	**315**
Accuracy	88.60	93.70	95.12	95.35	**95.81**	95.58	93.95
k-Nearest Neighbors classifier							
n_neighbors parameter	**2**	**5**	**7**	**10**	**11**	**12**	**13**	**15**	**20**
Accuracy	68.14	74.42	75.81	74.88	75.35	77.12	**77.44**	75.58	74.88
Support Vector classifier									
Kernel parameter	**linear**	**rbf**	**poly**	**sigmoid**					
Accuracy	**83.72**	83.49	81.69	81.40					
Logistic Regression classifier									
Solver parameter	**lbfgs**	**Newton-cg**	**liblinear**	**sag**	**saga**				
Accuracy	**84.42**	**84.42**	83.95	**84.42**	**84.42**				
XGB classifier									
n_estimators parameter	**2**	**3**	**4**	**6**	**8**	**10**			
Accuracy	94.42	94.65	95.35	**95.58**	95.35	95.35			
CatBoost classifier									
Learning rate parameter	**0.09**	**0.0009**	**0.0009**	**0.00009**					
Accuracy	96.51	96.51	**96.74**	96.51					

TABLE 8: Performance metric of accuracy, according to each classifier's main parameter tuning

Hyperparameter tuning played a critical role in optimizing the performance of each classifier. For the RF model (Table *8*), increasing the number of DTs (n_estimators) from 10 to 315 led to a consistent improvement in accuracy, culminating in the highest value of 95.81% at 300 trees. Beyond this point, performance plateaued and slightly declined, suggesting overfitting or diminishing returns. Similarly, XGBoost showed incremental performance gains with more estimators, achieving a peak accuracy of 95.58% at 6 n_estimators, beyond which no further improvement was observed. The CatBoost classifier exhibited high stability across different learning rates, with 96.74% accuracy at a learning rate of 0.0009.

In contrast, the KNN model (Table *8*) displayed more sensitivity to the n_neighbors parameter. Accuracy

peaked at 77.44% with 13 neighbors, showing a general upward trend from smaller neighbor counts, though with fluctuations, implying that model stability improves with increased neighborhood size but may plateau or regress due to excessive smoothing. For the SVC, the parameter of kernel had a relatively minor effect, with the linear kernel yielding the highest accuracy of 83.72%, followed closely by rbf, poly, and sigmoid options, suggesting the linear kernel best aligns with the structure of the data in this case.

Finally, LR (Table 8) demonstrated robustness across solvers, with lbfgs, Newton-cg, sag, and saga all achieving 84.42% accuracy, indicating convergence to similar optimal solutions. Overall, the hyperparameter tuning process confirms that methods like RF, XGBoost, and CatBoost are not only high-performing but also sensitive to appropriate parameter selection, while traditional models show greater stability but lower peak performance.

Furthermore, the accuracy curves of the MLP NN method are illustrated in Figure 3. As shown, this classifier achieves 85.58% accuracy after about 25 epochs.

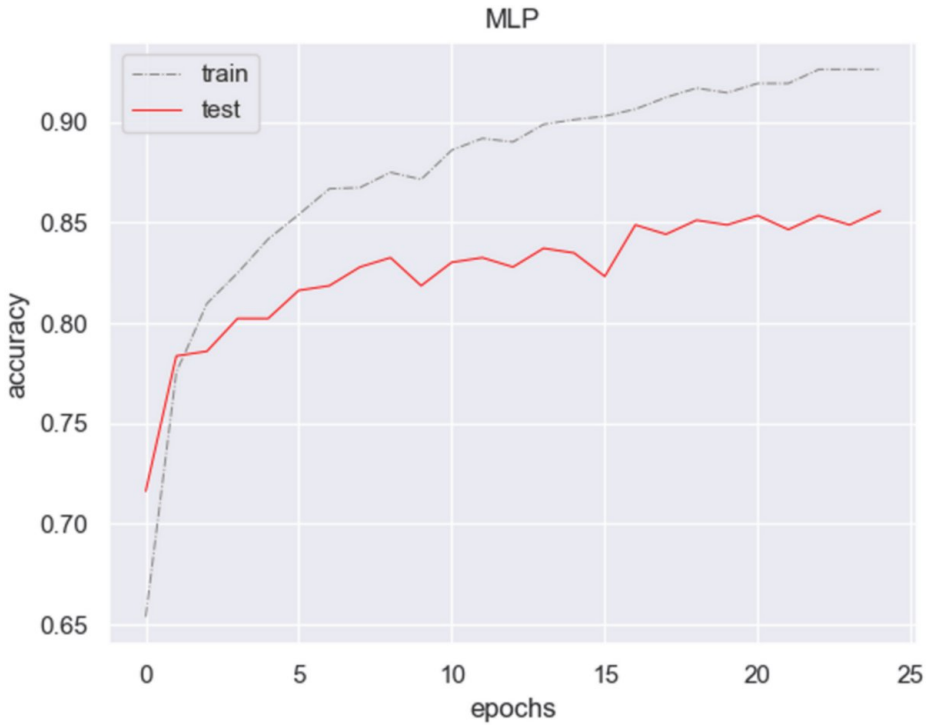

FIGURE 3: The accuracy curves of the multi-layer perceptron (MLP) neural network method

Performance of classifiers

The performance indices of all the classifiers we conducted are shown in Table 9.

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Specificity (%)	Cohen's Kappa (%)	ROC-AUC (%)
RF	95.81	96.77	92.02	94.34	**98.13**	91.02	95.08
KNN	77.44	85.11	49.08	62.26	94.76	47.78	71.92
SVC	83.72	82.52	72.39	77.12	90.64	64.57	81.51
LR	84.42	84.78	71.78	77.74	92.13	65.88	81.96
XGB	95.58	95.57	92.64	94.08	97.38	90.56	95.01
CatBoost	**96.74**	**96.86**	**94.48**	**95.65**	**98.13**	**93.05**	**96.30**
MLP	85.58	82.27	75.82	78.91	90.97	67.99	83.04

TABLE 9: Performance measures comparison of several classifiers

Bold values indicate the highest performance scores among all classifiers for each respective metric.

RF, Random Forest; KNN, K-Nearest Neighbors; SVC, Support Vector Classifier; LR, Logistic Regression; XGB, Extreme Gradient Boosting; MLP, Multi-Layer Perceptron; ROC, Receiver Operating Characteristic; AUC, Area Under the Curve

The performance evaluation of the classifiers developed in this study reveals that ensemble learning methods outperform traditional and neural classifiers across all key metrics, as demonstrated in Table *9*. The CatBoost model, which consists of an ensemble of decision trees, achieved the highest overall accuracy (96.74%), along with superior precision (96.86%), recall (94.48%), F1-score (95.65%), specificity (98.13%), Cohen's kappa (93.05%), and ROC-AUC (96.30%), indicating excellent predictive power and strong agreement with true class labels. The RF classifier closely followed, with an accuracy of 95.81% and an F1-score of 94.34%, further demonstrating the efficacy of ensemble techniques. The XGBoost model also performed comparably well, attaining an accuracy of 95.58%, F1-score of 94.08%, and ROC-AUC of 95.01%, reinforcing the consistency and robustness of gradient-boosted frameworks in classification tasks. In contrast, traditional classifiers such as LR and SVC showed moderate performance, with accuracies of 84.42% and 83.72%, respectively, and F1-scores around 77%, suggesting limitations in capturing complex, non-linear relationships in the data. MLP, as a representative of NNs, achieved 85.58% accuracy and an F1-score of 78.91%, outperforming traditional classifiers but still trailing behind ensemble methods. The KNN classifier demonstrated the weakest performance, with low recall (49.08%) and F1-score (62.26%) despite relatively high precision (85.11%) and specificity (94.76%). Overall, the findings affirm that ensemble-based models, particularly CatBoost and RF, provide the most reliable and generalizable solutions for the classification task in this study, exhibiting high discriminative capability and stability across all evaluated metrics.

Approach	Method	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Specificity (%)	Cohen's Kappa (%)	ROC-AUC (%)
Suñé [23]	Decision Tree Significant variants (Scaler: MinMax)	93.4884	-	-	-	-	-	-
Suñé [23]	RF Significant variants (Scaler: MinMax)	93.1783	-	-	-	-	-	-
Suñé [23]	Decision Tree (Scaler: MinMax)	92.2481	-	-	-	-	-	-
Suñé [23]	RF (Scaler: MinMax)	92.093	-	-	-	-	-	-
Suñé [23]	RF (Scaler: Standard)	91.6279	-	-	-	-	-	-
Suñé [23]	Decision tree (Scaler: Standard)	91.6279	-	-	-	-	-	-
Suñé [23]	SVM Significant variants (Scaler: MinMax)	88.9922	-	-	-	-	-	-
Suñé [23]	NN Significant variants (Scaler: Standard)	0.866667	-	-	-	-	-	-

Suñé [23]	KNN Significant variants (Scaler: Standard)	86.1679	-	-	-	-	-	-
Suñé [23]	LR Significant variants (Scaler: Standard)	82.3256	-	-	-	-	-	-
Suñé [23]	LR (Scaler: Standard)	81.7054	-	-	-	-	-	-
Suñé [23]	SVM (Scaler: MinMax)	81.7054	-	-	-	-	-	-
Suñé [23]	LR (Scaler: MinMax)	81.3953	-	-	-	-	-	-
Suñé [23]	NN (Scaler: Standard)	80.9302	-	-	-	-	-	-
Suñé [23]	SVM (Scaler: Standard)	80.9302	-	-	-	-	-	-
Suñé [23]	KNN (Scaler: Standard)	65.8244	-	-	-	-	-	-
Balbal [24]	XGBoost	95.35	-	-	-	-	-	-
Balbal [24]	RF	94.12	-	-	-	-	-	-
Balbal [24]	DT	91.33	-	-	-	-	-	-
Balbal [24]	AdaBoost	89.78	-	-	-	-	-	-
Balbal [24]	SVM	83.59	-	-	-	-	-	-
Balbal [24]	Gaussian Naive Bayes	76.47	-	-	-	-	-	-
Balbal [24]	KNN	75.54	-	-	-	-	-	-
Sendhil et al. [25]	RF	92.79	91 (Precision – 0) 96 (Precision – 1)	98 (Recall – 0) 0.83 (Recall – 1)	92 (Macro Average F1-score)	-	-	-
Sendhil et al. [25]	KNN	92.56	0.90 (Precision – 0) 0.98 (Precision – 1)	99 (Recall – 0) 82 (Recall – 1)	92 (Macro Average F1-score)	-	-	-
Sendhil et al. [25]	SVM	49.30	0.46 (Precision – 0) 52 (Precision – 1)	47 (Recall – 0) 52 (Recall – 1)	49 (Macro Average F1-score)			
Sendhil et al. [25]	LR	48.14	45 (Precision – 0) 51 (Precision – 1)	45 (Recall – 0) 51 (Recall – 1)	48 (Macro Average F1-score)			
Sendhil et al. [25]	GBM	45.35	42 (Precision – 0) 49 (Precision – 1)	40 (Recall – 0) 50 (Recall – 1)	45 (Macro Average F1-score)	-	-	-
Airlangga [26]	MLP	84.41	84.30	84.41	84.17	-	-	-
Airlangga [26]	Convolutional NN	0.8865	88.84	88.65	88.62	-	-	-
Airlangga [26]	Long Short-Term Memory	0.7557	75.67	75.57	75.28	-	-	-
Raghupathy et al. [27]	GB (with Shapley Additive Explanations)	96.91	91.70	94.10	92.90	91.66	-	-
Raghupathy et al. [27]	LightGBM (with Shapley Additive Explanations)	95.11	92.80	93.40	93.10	92.80	-	-
Raghupathy et al. [27]	XGBoost (with Shapley Additive Explanations)	94.88	91.60	94.70	92.90	91.13	-	-
Raghupathy et al. [27]	LightGBM (without Shapley Additive Explanations)	94.88	92.20	93.40	92.80	92.20	-	-
Raghupathy	XGBoost (without Shapley Additive	94.65	91.10	93.40	92.50	91.61	-	-

et al. [27]	Explanations)							
Raghupathy et al. [27]	GB (without Shapley Additive Explanations)	94.64	91.10	94.10	92.60	91.08	-	-
Oncu [28]	Artificial NN	87.08	88	87	87	-	-	-

TABLE 10: Competitors' performance using the dataset

RF, Random Forest; SVM, Support Vector Machine; NN, Neural Network; KNN, K-Nearest Neighbors; LR, Logistic Regression; XGBoost, Extreme Gradient Boosting; DT, Decision Tree; GBM, Gradient Boosting Machine; MLP, Multi-Layer Perceptron; GB, Gradient Boosting

Table *10* presents a comparative analysis of several ML models applied to the same dataset [22] as our study, showcasing the performance metrics reported by various studies. The empty cells in the table indicate performance fields where referenced studies do not cover. The models proposed in this work consistently outperform existing approaches across multiple evaluation criteria. Raghupathy et al. [27] reported the highest accuracy of 96.91% using GB by implementing data cleaning with shapley additive explanations values; however, their F1-score (92.90%) and lack of ROC-AUC reporting limit comprehensive comparison. Balbal [24] and Suñé [23] also explored ensemble methods such as XGBoost and RF, achieving accuracies up to 95.35% and 93.49%, respectively, but did not report other critical metrics such as precision, recall, or F1-score, hindering a thorough performance assessment. Among neural models, Airlangga [26] found convolutional NNs to outperform MLP and long short-term memory architectures, with an accuracy of 88.65%, which remains below the performance of this study's ensemble models. Notably, the models by Sendhil et al. [25] demonstrated inconsistent results, with RF achieving 92.79% accuracy and a macro-average F1-score of 92%, whereas SVM, LR, and GBM models performed poorly, with accuracies below 50%. Overall, we observe from Table *10* that all GBM models listed in four similar publications showed significant performance indicators, and only in one of them did they lag behind. Additionally, the comparison results indicate that the models developed in this work, particularly CatBoost and RF, offer state-of-the-art performance, not only in accuracy but also in balanced classification metrics such as precision, recall, F1-score, specificity, and ROC-AUC, thus establishing their robustness and generalizability for the target prediction task.

Overall, by comparing the performance measures of all classifiers we implemented, we deduce that the CatBoost classifier outperforms all the models, with the only exception of RF that achieves an equalizer in the specificity index.

Conclusions

Estimating the probability of AD based on patients' medical and demographic information proved to be a very powerful tool. The machine learning techniques we applied to predict AD demonstrated high performance indicators. The CatBoost model achieved the highest overall accuracy (96.74%), along with superior precision (96.86%), recall (94.48%), F1-score (95.65%), specificity (98.13%), Cohen's kappa (93.05%), and ROC-AUC (96.30%), indicating excellent predictive power and strong agreement with true class labels. The RF classifier closely followed, with an accuracy of 95.81% and an F1-score of 94.34%, further demonstrating the efficacy of ensemble techniques. The XGBoost model also performed comparably well, attaining an accuracy of 95.58%, F1-score of 94.08%, and ROC-AUC of 95.01%, reinforcing the consistency and robustness of gradient-boosted frameworks in classification tasks. In contrast, traditional classifiers such as LR and SVC showed moderate performance, with accuracies of 84.42% and 83.72%, respectively, and F1-scores around 77%, suggesting limitations in capturing complex, non-linear relationships in the data. MLP, as a representative of NNs, achieved 85.58% accuracy and an F1-score of 78.91%, outperforming traditional classifiers but still trailing behind ensemble methods. The KNN classifier demonstrated the weakest performance, with low recall (49.08%) and F1-score (62.26%) despite relatively high precision (85.11%) and specificity (94.76%), indicating a strong tendency toward false negatives and poor balance. Overall, the findings affirm that the models CatBoost and RF provide the most reliable and generalizable solutions for the classification task in this study, exhibiting high discriminative capability and stability across all evaluated metrics. It is finally concluded that the integration of machine learning mechanisms in medicine provides an additional tool in the early diagnosis of AD and, thus, contributes to the improvement of medical treatment or lifestyle interventions. In future work, we plan to study the application of ML to other serious diseases. Furthermore, the machine learning diagnostic tool of our study could be integrated in the future into a web-based server application, where clinicians could operate the tool and generate diagnoses, through a suitable user interface.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of

the work.

Concept and design: Evgenia Psarra, Dimitris Apostolou

Acquisition, analysis, or interpretation of data: Evgenia Psarra, Dimitris Apostolou

Drafting of the manuscript: Evgenia Psarra, Dimitris Apostolou

Critical review of the manuscript for important intellectual content: Evgenia Psarra, Dimitris Apostolou

Supervision: Evgenia Psarra, Dimitris Apostolou

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Kaur I, Sachdeva R: Prediction models for early detection of Alzheimer: recent trends and future prospects. Archives of Computational Methods in Engineering. 2025, 32:3565-3592. 10.1007/s11831-025-10246-3
2. Rony MKK, Parvin MR, Wahiduzzaman M, Akter K, Ullah M: Challenges and advancements in the health-related quality of life of older people. Advances in Public Health. 2024, 2024:8839631. 10.1155/2024/8839631
3. Albar SA: Factors influencing omega-3 knowledge and consumption among Saudi adults: a cross-sectional study. Advances in Public Health. 2025, 2025:9976679. 10.1155/adph/9976679
4. Sampat NA, AlRizeiqi M, Mohamed Essa M, Al-Subhi L, Al-Adawi S, Qoronfleh MW: Assessing general knowledge and nutritional knowledge on the role of vitamin D, attitudes, and behavior toward sun exposure among tertiary students in Oman. Advances in Public Health. 2024, 2024:9544155. 10.1155/2024/9544155
5. Rashed M, Hossain MI, Mahdi A, Mustofa G: Hybrid machine learning models for accurate type 2 diabetes mellitus prediction using a stacking classifier and a meta-model approach. Cureus Journal of Computer Science. 2025, 1-23. 10.7759/s44389-025-03135-0
6. Rawat P, Singh PD, Tripathi V: A fast and reliable transformer-based TabPFN model for liver disease diagnosis. Cureus Journal of Computer Science. 2025, 1-24. 10.7759/s44389-025-04072-8
7. Rajpoot NK, Singh PD, Pant B, Sharma M: An intelligent framework for real-time heart stroke monitoring and early detection using machine learning. Cureus Journal of Computer Science. 2025, 1-20. 10.7759/s44389-025-04004-6
8. Priya B, Gupta P, Singh S: Harnessing the potentials of machine learning models in Alzheimer's disease prediction and detection. Advances in Computational Methods and Modeling for Science and Engineering. 2025, 259-267. 10.1016/B978-0-44-330012-7.00027-8
9. Mao S, Zhao X, Wang L, Man Y, Li K: Palmitoylation-related gene ZDHHC22 as a potential diagnostic and immunomodulatory target in Alzheimer's disease: insights from machine learning analyses and WGCNA. European Journal of Medical Research. 2025, 30:46. 10.1186/s40001-025-02277-0
10. Jiao B, Ouyang Z, Xiao X, et al.: Development and validation of machine learning models with blood-based digital biomarkers for Alzheimer's disease diagnosis: a multicohort diagnostic study. eClinicalMedicine. 2025, 81:103142. 10.1016/j.eclinm.2025.103142
11. Javed R, Abbas T, Shahzad T, Kanwal K, Ramay SA, Khan MA, Ouahada K: Enhancing chronic disease prediction in IoMT-enabled healthcare 5.0 using deep machine learning: Alzheimer's disease as a case study. IEEE Access. 2025, 13:14252-14272. 10.1109/access.2025.3525514
12. Psarra E, Verginadis Y, Patiniotakis I, Apostolou D, Mentzas G: A context-aware security model for a combination of attribute-based access control and attribute-based encryption in the healthcare domain. Web, Artificial Intelligence and Network Applications. Barolli L, Amato F, Moscato F, Enokido T, Takizawa M (ed): Springer, Cham; 2020. 1150:1133-1142. 10.1007/978-3-030-44038-1_104
13. Psarra E, Patiniotakis I, Verginadis Y, Apostolou D, Mentzas G: Securing access to healthcare data with context-aware policies. 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA), Piraeus, Greece. 2020, 1-6. 10.1109/IISA50023.2020.9284393
14. Psarra E, Verginadis Y, Patiniotakis I, Apostolou D, Mentzas G: Accessing electronic health records in critical incidents using context-aware attribute-based access control. Intelligent Decision Technologies. 2021, 15:667-679. 10.3233/idt-210214
15. Psarra E, Apostolou D, Verginadis Y, Patiniotakis I, Mentzas G: Context-based, predictive access control to electronic health records. Electronics. 2022, 11:3040. 10.3390/electronics11193040
16. Psarra E: Clinical Decision Support Systems - Diagnostic and Prognostic Attribute-Based Access Control in Acute Care (Doctoral Thesis). University of Piraeus, Pireas, Greece; 2023. 10.26267/unipi_dione/3419
17. Psarra E, Apostolou D, Verginadis Y, Patiniotakis I, Mentzas G: Permissioned blockchain network for proactive access control to electronic health records. BMC Medical Informatics and Decision Making. 2024, 24:303. 10.1186/s12911-024-02708-8

18. de Oliveira MT, Verginadis Y, Reis LHA, Psarra E, Patiniotakis I, Olabarriaga SD: AC-ABAC: Attribute-based access control for electronic medical records during acute care. Expert Systems with Applications. 2023, 213:119271. 10.1016/j.eswa.2022.119271
19. Psarra E, Apostolou D: Dynamic and personalized access control to electronic health records. Advances in Artificial Intelligence-Empowered Decision Support Systems. Tsihrintzis GA, Virvou M, Doukas H, Jain LC (ed): Springer, Cham; 2024. 39:129-153. 10.1007/978-3-031-62316-5_5
20. Psarra E, Apostolou D: A multi-layer perceptron neural network-based access control mechanism to breast cancer patients' health data. 2024 15th International Conference on Information, Intelligence, Systems & Applications (IISA), Chania Crete, Greece. 2024, 1-4. 10.1109/IISA62523.2024.10786721
21. Psarra E, Apostolou D: A neural network-based context handling mechanism for accessing the medical information of patients suffering from coronary heart disease. 2024 15th International Conference on Information, Intelligence, Systems & Applications (IISA), Chania Crete, Greece. 2024, 1-6. 10.1109/IISA62523.2024.10786708
22. Kaggle: Alzheimer's disease prediction. (2024). Accessed: January 1, 2025: https://www.kaggle.com/code/adhamtarek147/alzheimer-s-disease-prediction.
23. Suñé LS: Machine Learning techniques for Alzheimer prediction. Statistical analysis of Alzheimer data. (2025). Accessed: January 1, 2025: http://hdl.handle.net/10609/152036.
24. Balbal KF: Machine learning approaches for prediction of Alzheimer's disease. Journal of Artificial Intelligence and Data Science. 2024, 4:110-118.
25. Sendhil R, Patibandha N, Shah P, Mandal A: Comparative evaluation of machine learning techniques for early Alzheimer's detection: challenges and insights. 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI), Erode, India. 2025, 1050-1058. 10.1109/ICMSCI62561.2025.10893998
26. Airlangga G: Advancing Alzheimer's diagnosis: a comparative analysis of deep learning architectures on multidimensional health data. Jurnal Informatika Ekonomi Bisnis. 2024, 6:810-814. 10.37034/infeb.v6i4.1046
27. Raghupathy BK, Reddy MR, Theeda P, Balasubramanian E, Namachivayam RK, Ganesan M: Harnessing explainable artificial intelligence (XAI) based SHAPLEY values and ensemble techniques for accurate Alzheimer's disease diagnosis. Engineering, Technology & Applied Science Research. 2025, 15:20743-20747. 10.48084/etasr.9619
28. Oncu E: Leveraging AI models for detection and classification of Alzheimer's disease using MRI and clinical data [PREPRINT]. Preprints. 2025, 1-16. 10.20944/preprints202501.0061.v1