Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 6 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i06.60489

PAPER

MindCare: Enabling Accessible Healthcare Through a Web-Based Predictive Tool for Alzheimer's Disease Diagnosis

Aunsia Khan[1], **Khausaalyaah Sinathurai**[1], **Anusha Achuthan**[1] (✉), **Anton Lord**[2]

[1]School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia

[2]Metro South Addiction and Mental Health Services, Queensland, Australia

anusha@usm.my

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disease and one of the leading causes of cognitive decline in the older population. Conventional diagnostic tools often rely on clinical symptoms that appear in later stages, highlighting the need for an accurate and accessible tool for early prediction. This study introduces *MindCare*, a web-based application for AD prediction using clinical and neuropsychological data. Multiple machine learning algorithms were implemented to generate predictive models, evaluate their performance, and highlight important features that contribute to classification. This application provides real-time predictions for new patient instances, comparative visualization of classifier performance, and correlation analysis of features with the outcome. Furthermore, an administrator module supports model training, continuous updating, and performance monitoring to ensure reliability over time. Decision Tree (DT) produced the most effective results, achieving an accuracy of 98% with an average F1 score of 95.7%. The findings demonstrate that integrating machine learning into a decision-support system can provide accurate and interpretable predictions of AD. *MindCare* offers a practical, user-friendly tool that combines predictive analytics with clinical insight, highlighting its potential to predict early and contribute to improved care strategies for individuals at risk of Alzheimer's disease.

KEYWORDS

Alzheimer's disease (AD) prediction system, sustainable development goal, machine learning algorithms, early prediction, feature selection

1 INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disease that leads to memory impairment, cognitive decline, and finally loss of independence in old people. It is becoming an increasingly urgent public health problem, particularly

Khan, A., Sinathurai, K., Achuthan, A., Lord, A. (2026). MindCare: Enabling Accessible Healthcare Through a Web-Based Predictive Tool for Alzheimer's Disease Diagnosis. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(6), pp. 172–185. https://doi.org/10.3991/ijoe.v22i06.60489

Article submitted 2026-01-05. Revision uploaded 2026-02-28. Final acceptance 2026-03-03.

as the world population ages. The identification of the disease in early stages is important since it allows timely interventions that may delay disease progression and quality of life for elderly patients [1, 2]. However, AD is difficult to recognize at an early stage since the first symptoms are often overlooked or misdiagnosed as normal aging [3, 4]. Currently, imaging techniques and invasive biomarkers are widely used for diagnosis, which are costly and inaccessible, especially in resource-limited settings. Machine learning has become a great tool for the education [5] and healthcare industries and in helping the detection and prediction of AD. It facilitates early diagnosis and risk assessment by analyzing complicated patterns in various data sets [6–9]. Previous research studies used various machine learning models, such as support vector machines (SVM), Naive Bayes (NB), and logistic regression, for AD prediction. These models have proved to be more accurate and reliable, especially when used in combination with biomarkers or medical imaging data.

Recent researches has shown a growing trend of using artificial intelligence based tools through web and mobile platforms [10]. Many of these systems are focused on the neuroimaging data and deep learning techniques. For instance, [11] and [12] developed a web-based tool that accepts structural MRI as input and a convolutional neural network (CNN) for the prediction of AD. The systems achieved high predictive performance and provided visualization capabilities. However, the imaging-based approaches require specialized neuroimaging infrastructures, structured images, and high computational resources. This limits their applicability, particularly in hospitals with limited resources.

Other studies have explored lightweight screen tools deployed in web or mobile environments. A study proposed a web-based tool for preclinical cognitive screening [13], while [14] introduced a predictive mobile/web application using ML models for AD prediction. Similarly, the Streamlit-based system by [15] demonstrated ML models such as logistic regression in an accessible interface for AD prediction. Extended digital screening approaches were later proposed using ML models with handwriting analysis [16]. These applications emphasize accessibility and scalability; however, they either rely on structured datasets or specific inputs with limited predictive pipelines or do not fully integrate clinical health records into their workflow. In contrast to ML models, [17] highlighted the importance of explainable AI by integrating multimodal clinical and imaging data into interpretable ML models. Meanwhile, [18] developed a web-based tool to estimate the burden of AD using projections. Although these works are significant contributions to interpretability and planning in public health, they do not directly address real-time predictive modeling using clinical health records. Moreover, there are not many ML-based diagnostic tools available that provide clinician-friendly platforms (refer to Table 1). Furthermore, the lack of intuitive interfaces with feedback and clear visual representations further hinders their implementation in the real world [19]. The existing systems use a single machine learning algorithm within the frameworks for comparative analysis. Clinicians are struggling to identify the optimal prediction tool, since there is no good way to assess the effectiveness of various models [20]. These challenges leave a gap for an integrated, accessible solution that brings together machine learning techniques and real-world clinical applications. This research is based on the increasing need for more effective, trustworthy, and affordable tools for AD prediction [21–23]. However, traditional diagnostic methods find it difficult to detect Alzheimer's in its initial stages, thus missing opportunities for treatment that may be more successful if started at an earlier stage.

Table 1. Some of the existing mobile or web-based applications for different types of data and machine learning algorithms

Ref.	Data Type	Platform	ML/Algorithm	Evaluation Metrics	Visualization	Key Limitation
[11]	MRI scans	Web	Deep Learning (VGG-16 CNN)	Accuracy	Visualisation of MRI	Requires MRI scans, high computation, and limited real-world hospital integration
[12]	MRI scans	Web	Deep Learning (CNN pipelines)	Accuracy	Visualisation of MRI	Requires imaging infrastructure; does not integrate routine clinical records
[13]	Cognitive	Web	Not mentioned	Not mentioned	results visualization	Limited to task-based data; small-scale evaluation
[14]	Cognitive + basic demographic	Mobile/ Web	SVM, NB	Accuracy	Output only	Focus on a small dataset, limited feature exploration, and minimal visualisation
[15]	age, cognitive scores, genetic	Web (Streamlit)	Logistic Regression	Accuracy (~70%)	Output only	Single classifier, no feature importance or comparative analysis
[16]	Handwriting	Mobile/ Web	Supervised ML	Accuracy (~88%)	Output only	Task-specific input; lacks hospital EMR integration
[17]	MRI + clinical + psych	Web	Random Forest, Explainable AI	Accuracy (98.8%), SHAP explanations	Feature contribution	Complex multimodal data; not fully hospital-ready; high dimensionality
[18]	Epidemiological data	Web	Simulation/ Projection	Projection outputs	Projection plots	Not individual-level prediction; not ML-based

The use of daily clinical health records, lack of transparency, and variable accuracy across multimodal datasets persist. Clinical health records contain rich demographic and cognitive test information that can support early prediction without requiring highly invasive imaging data. However, effectively utilizing clinical health data requires preprocessing, feature selection, model comparison, and mechanisms to ensure reliability and clinical interpretability.

To address this, the present work proposes a web-based tool that operates directly on clinical health records, including demographic and neuropathological data. The objective is to improve the decision-making process for AD by incorporating a predictive system into an intuitive web application. The system includes an analytical pipeline comprising preprocessing and feature selection, followed by classification using machine learning algorithms such as SVM, NB, and DT. Unlike other existing platforms that report single-model accuracy, the proposed system provides a comparative performance evaluation using accuracy, precision, recall, and F1 score. The proposed system also provides visualization of selected features and classification outcomes. By integrating the analysis and prediction module with an interactive web interface, this application aims to bridge the gap between ML and decision support systems.

2 MATERIALS AND METHODS

This study presents a web-based server ("MindCare") for AD risk assessment with real-time feedback. MindCare can be accessed via a computer system or a mobile phone and integrated into a healthcare system. It comprises three primary modules, namely (1) Data Entry and Management, (2) Machine Learning Assessment, and (3) Results Visualization. Each module plays an important role in the overall workflow. It provides predictions using supervised classification machine learning algorithms, including NB, SVM and DT, to help healthcare practitioners analyze the risk factors for AD. This web application offers a single platform for data entry, predictive analysis, and result visualization (see Figure 1). Following this sequence,

clinical data is collected and fed to the preprocessing block. Unlike many existing web-based tools, which assume clean, ready-to-use datasets, the proposed system includes preprocessing to handle missing values, normalize and scale the data for the next step. Key features are selected after hot encoding, making the data ready for analysis. Feature selection serves two important purposes:

1. Improves computational efficiency
2. Identifies clinically relevant features

The inclusion of this module distinguishes the system from several existing web-based systems that directly apply classifiers without dimensionality reduction or feature selection.

In the third and fourth steps, the data is split into testing, training, and validation sets. The data is analyzed using DT, NB, and SVM. These algorithms were selected for their suitability for structured tabular data, computational efficiency, and interpretability. Unlike deep learning models that require high-dimensional data and extensive computational resources, these methods are well aligned with moderate-sized clinical datasets. Performance is evaluated using accuracy, recall, precision, and F1 score. The best model is selected; its visualization and the outcome of the risk prediction are plotted, along with the relationship between selected features and the target class. This entire process occurs at the back end of the application. The backend (developed using the Flask web framework for Python) responds to the dashboard of the frontend application for clinicians (developed using a modern web stack including React, TypeScript, Vue, and Tailwind CSS).

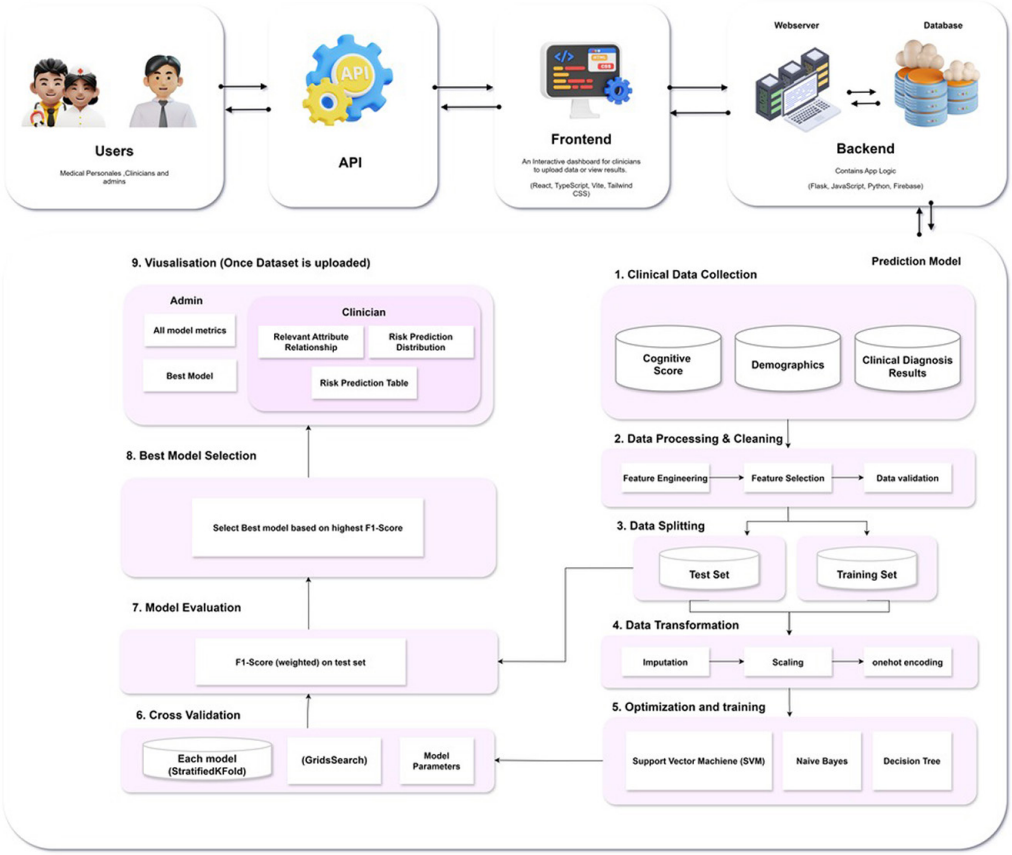

Fig. 1. System architecture diagram for MindCare

2.1 Data entry and management module

The Data Entry and Management Module governs the authentication workflow, data input, input validation, and the preprocessing pipeline (see Figure 2). User access to the system is managed via Firebase authentication, which supports both email/password registration and login. Role-based access control (RBAC) is enforced by conditions that enable features based on user metadata stored in Firebase. Clinicians can register directly via the front end, while admin accounts are pre-registered manually through Firebase's admin console to prevent unauthorized access to critical operations such as model training. Data can be entered manually into the system or uploaded as a .csv file.

An important aspect of this pipeline is feature engineering, the process of creating features from raw data. For example, the participant's BIRTHYEAR is recorded; in this pipeline, the participant's auto-age is calculated. After feature engineering, the pipeline moves on to the feature selection and data cleaning stage. This involves the use of principal component analysis (PCA) to select features from the original data, resulting in the top twelve features found to be used in predicting the actual outcome (see Figure 3).

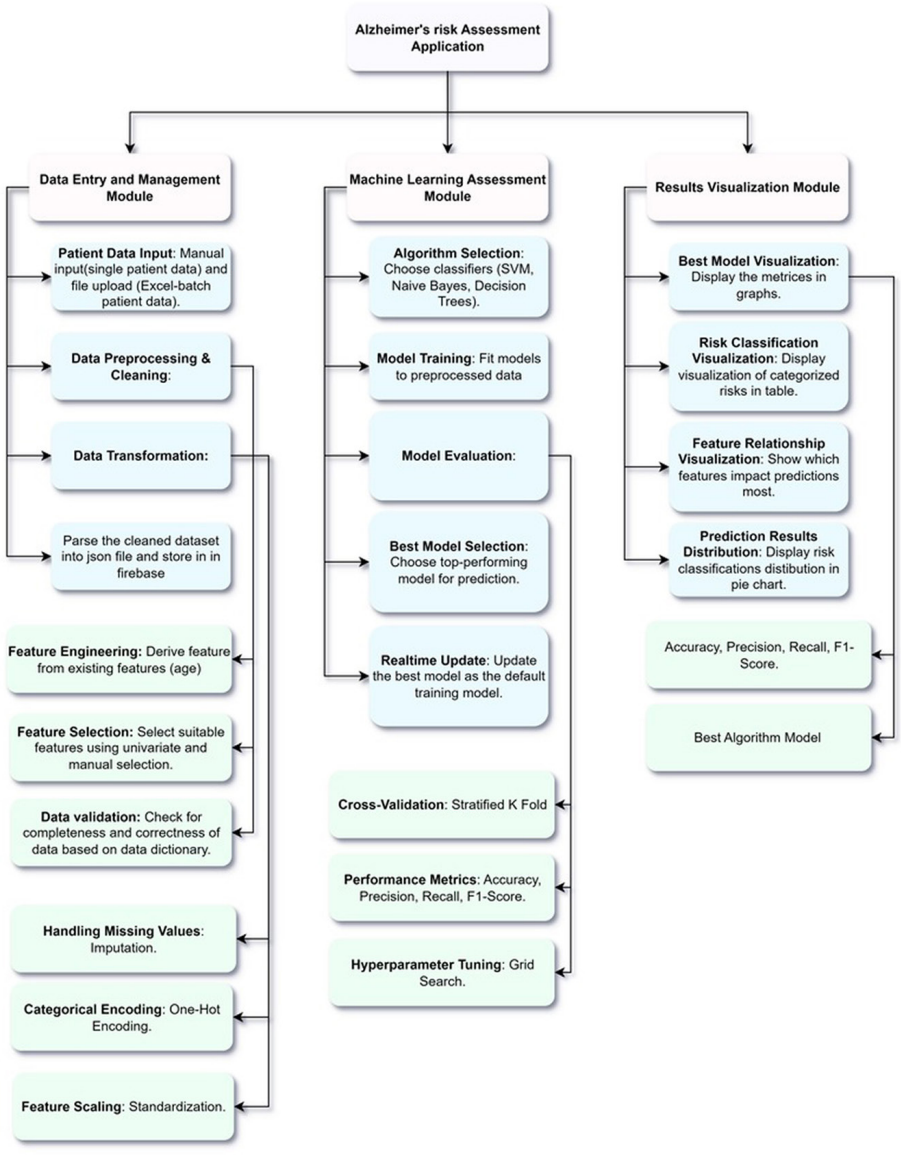

Fig. 2. Details of all the modules included in the MindCare application system

The final list of selected features includes demographic variables, cognitive tests, and clinical assessment scores that are frequently linked to cognitive decline and the progression of Alzheimer's. In the next step, missing values will be handled, and any incomplete or unusable values will be discarded if necessary. For numerical features, missing values are imputed with the mean. For categorical features, missing values are imputed with the most frequent value. These imputations are handled using the 'SimpleImputer' from Scikit-learn, so no rows are dropped unnecessarily, and the data remains consistent. Through data validation, feature engineering, imputation, scaling, and encoding, the system will ensure reliable and accurate predictions across a diverse population.

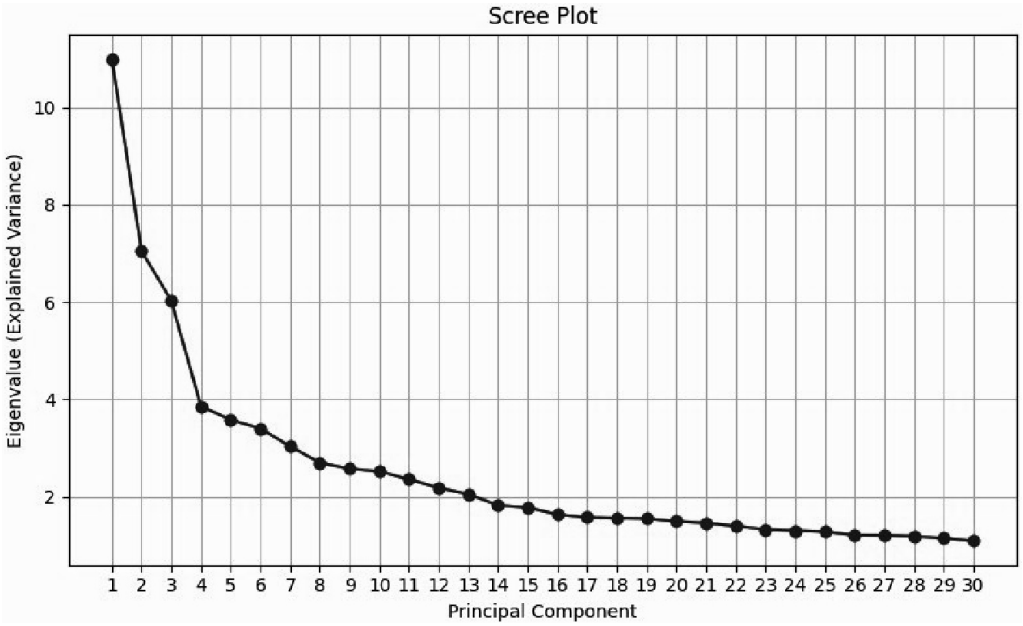

Fig. 3. Feature selection using PCA, emphasising the top twelve features for analysis

2.2 Machine learning assessment module

The Machine Learning Assessment Module includes the training, evaluation, and deployment of machine learning models directly from the web interface. It is accessible only to admin users via the Model Training tab in the application. It applies to structured datasets containing both input features and a labeled outcome. With preprocessing, the dataset is split into training and testing sets at an 80:20 ratio. The application supports three classification algorithms, using baseline hyperparameters and then optimized using 'GridSearchCV' with a 3-fold stratified cross-validation strategy. This not only tunes the hyperparameters but also addresses overfitting by validating the models with different data rates. Once training is complete, the model with the highest F1-score is automatically selected as the best model. This model is then serialized and saved using the JOBLIB library in a given backend directory. Alongside the model file, a separate file is stored to record the performance metrics for each trained model. This enables later review and transparency.

2.3 Results visualization module

The visualization module provides real-time, interactive data visualizations for the clinician and administrator to interpret predictions and assess model performance (see Figure 4). For example, patients with medical illnesses were more often placed in the impaired category, and distribution charts by sex again showed that cognition is more impaired in females. This module is available after predictions are generated using the "Generate Prediction" button in both the Manual Input and the Dataset Upload Workflows.

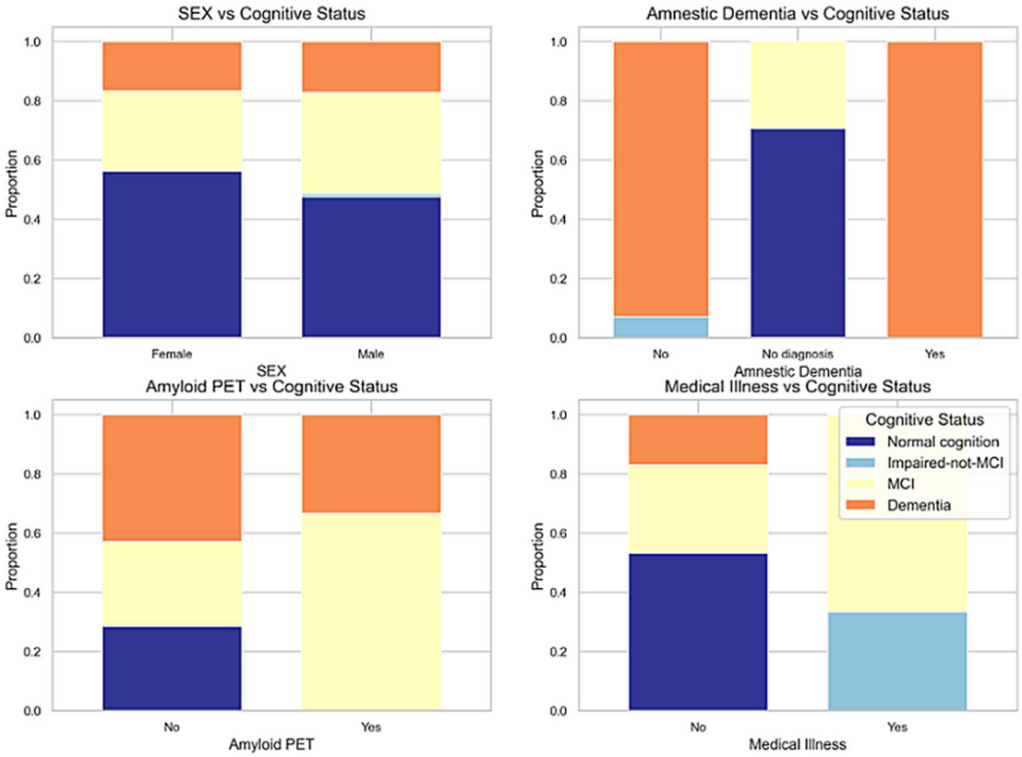

Fig. 4. Bar graph visualisation of the analysis of clinical features against the target class

Furthermore, another set of visuals is provided in the Model Training Results panel. After training, a comparative chart is presented that shows the performance metrics of all three models (see Figure 5). These are plotted using Matplotlib's subplot API and color-coded for easy visualization. This allows quick visual evaluation of trade-offs among models and helps in reasoning about the best choice of model.

2.4 Dataset

The dataset for this research was obtained from the National Alzheimer's Coordination Center and (NACC) frozen between 2005 and 2024 [24]. It comprises clinical data as well as neuropsychological data of participants with normal cognition, mild cognitive impairment, and those with AD (refer to Table 2). The full dataset of 750 participants is randomly sampled into 250 instances for each of the two test samples, Test 1 and Test 2, and a validation set, Test 3.

Fig. 5. Result visualisation for the full training set using the Visualization Module

2.5 Technology deployed

This section defines the hardware and software specifications, including the programming languages and algorithms used during the development of the MindCare application. The proposed application was assessed on a RedmiBook 16 with AMD Ryzen 7 4700U, Radeon Graphics, 16GB RAM, and Windows 11. The software used to run the web application is summarized in Table 3.

Table 2. Detailed description of the selected features and their value ranges

Feature	Short Description	Allowable Codes
NACCID	Subject ID number	Prefix "NACC" followed by 6 numerals
SEX	Subject's sex	1 = Male, 2 = Female
EDUC	Years of education	0 – 36, 99 = Unknown
AGE	Subject's age	

(Continued)

Table 2. Detailed description of the selected features and their value ranges *(Continued)*

Feature	Short Description	Allowable Codes
DYSILL	Presumptive etiologic diagnosis of the cognitive disorder – Cognitive impairment due to systemic disease/medical illness	0 = No (assumed assessed and found not present), 1 = Yes
NACCUDSD	Cognitive status at UDS visit	1 = Normal cognition, 2 = Impaired-not-MCI, 3 = MCI, 4 = AD
DYSILLIF	Primary, contributing, or non-contributing cause of cognitive impairment – systemic disease/medical illness	1 = Primary, 2 = Contributing, 3 = Non-contributing, 7 = Cognitively impaired but no diagnosis of impairment due to systemic disease/medical illness, 8 = Diagnosis of normal cognition, –4 = Not applicable
AMYLPET	Abnormally elevated amyloid on PET	0 = No, 1 = Yes, 8 = Unknown/not assessed, –4 = Not applicable
NACCPPAG	Dementia syndrome – Primary progressive aphasia (PPA) subtype according to the criteria outlined by Gorno-Tempini et al. 2011	1 = Meets criteria for semantic PPA, 2 = Meets criteria for logopenic PPA, 3 = Meets criteria for nonfluent/agrammatic PPA, 4 = PPA other/not otherwise specified, 7 = Impaired but no PPA syndrome, 8 = No cognitive impairment, –4 = Not applicable
AMNDEM	Dementia syndrome – Amnestic multidomain dementia syndrome	0 = No, 1 = Yes, 8 = No diagnosis of dementia, –4 = Not applicable
MOCATRAI	MoCA: Visuospatial/executive – Trails	0–1, 95 = Physical problem, 96 = Cognitive/behavior problem, 97 = Other problem, 98 = Verbal refusal, 4 = Not available
UDSBENTC	Total score for a copy of the Benson figure	0–17, 95 = Physical problem, 96 = Cognitive/behavior problem, 97 = Other problem, 98 = Verbal refusal, –4 = Not available

Table 3. Software used in the MindCare application

Software/Tool	Purpose	Description
Python	Backend development & ML processing	Core programming language used for backend logic data handling.
Flask	Web server & backend framework	Serves REST APIs and handles request routing for the web application.
JOBLIB	Model persistence	Serializes trained models and metrics for reuse in predictions.
matplotlib	Visualization generation	Used for plotting graphs such as performance metrics and risk distributions.
Firebase Authentication	User login and RBAC	Manages secure login for clinicians and admins, enforcing role-based access.
Firebase Realtime Database	Data storage	Stores uploaded datasets, prediction results, and metadata in a real-time JSON structure.
React.js	Frontend framework	Powers the dynamic and responsive user interface with reusable components.
Vite	Frontend development server & bundler	A fast build tool used to serve the React app during development and bundle it for production.
Tailwind CSS	UI styling framework	A utility-first CSS framework that simplifies styling React components with a clean, responsive design.
HTML/CSS & JavaScript	UI structure and behavior	Base technologies that support the frontend layout and interactivity.
DOTENV	Configuration loader	Manages environment variables for secure and flexible configuration (e.g., Firebase API keys, model paths).
VS Code	Code editing and running	IDE used to write and manage both frontend and backend code efficiently.

This study proposes a new robust early AD prediction system based on machine learning and web. The system provides a customized access framework for multiple clinicians. By employing role-based access control (RBAC) via Firebase authentication, the application distinguishes between clinicians and administrators.

Clinicians can focus on making predictions and analyzing results, while administrators gain access to advanced features, including model training and performance evaluation. This division not only provides better security but also aligns with clinical workflow practices in healthcare organizations. Furthermore, the ability to upload batches of datasets or enter data manually per patient is a highly flexible feature with significant potential for clinical use. By providing comparative graphs that show how each model performs, the platform encourages transparency and further improvement of the models.

3 RESULTS

The test sample, Test 1, was a stratified sample with 80% for training and 20% for testing. The predicted class distribution was very similar to the original dataset composition. Specifically, the model predicted 52.6% of cases as Normal Cognition (NC), 16.9% as AD, 30.1% as Mild Cognitive Impairment (either amnestic or non-amnestic) (MCI), and only 0.4% as Cognitively Impaired but not MCI (NMCI). This distribution matches well with the proportions of the original dataset, and the model generalized well to unseen data.

Test 2 retested the same process on a different set of data to prove the model's consistency. The predictions showed small changes: 59.2% NC, 6.8% AD and 33.6% MCI. 0.3% was predicted as NMCI. The increase in the number of NC and MCI predictions and the reduction in the number of AD, predictions indicate minor variance while maintaining class balance. These results emphasize that the trained model gives stable predictions across different test splits and does not change the original dataset too much.

The process of tuning and evaluating the model included comparing three supervised learning algorithms, namely SVM, NB, and DT. Hyperparameter tuning was performed using 3-fold cross-validation, making it robust. Each model was tested across several parameter configurations, with the best combination producing the highest F1 score on the training data. Once the tuning phase was complete, all models were evaluated on the test set to determine their generalization performance. The SVM achieved a weighted F1 score of 0.9486, NB 0.9438, and the DT performed best with a test F1 score of 0.9835 (refer Table 4). Note also that DT not only sustained significant training performance but also showed excellent generalization, as indicated by the small drop in test performance relative to training performance.

4 DISCUSSION

A class imbalance was reflected in Tests 1 and 2 by using a full data set. The stratified sampling technique used for test splitting helps maintain the class proportion, but the model stability depends upon a larger volume and diversity of the training dataset. It tends to make more stable and generalizable predictions with consistent F1 scores (refer to Table 4). Based on these results, DT was the best model, which offered an F1score of 98.35%. It provided a balanced combination of interpretability, accuracy, and efficiency. The correlation between the training results and the test results makes it the most suitable candidate to be deployed inside the MindCare application.

Table 4. Overall result for each classifier with the full dataset and different test samples

	Accuracy	Precision	Recall	F1-Score
Decision Tree Classifier				
Full Training set	98.67%	98.74%	98.67%	98.35%
Test 1	94.00%	92.33%	94.00%	93.13%
Test 2	98.00%	96.14%	98.00%	97.04%
Test 3	96.00%	96.00%	96.00%	96.00%
Naïve Bayes Classifier				
Full Training set	78.67%	98.17%	78.67%	79.52%
Test 1	80.00%	96.67%	80.00%	78.81%
Test 2	90.00%	95.40%	90.00%	92.35%
Test 3	98.00%	96.17%	98.00%	97.04%
SVM Classifier				
Full Training set	92.67%	90.78%	92.67%	91.69%
Test 1	88.00%	84.59%	88.00%	85.93%
Test 2	88.00%	86.38%	88.00%	86.25%
Test 3	82.00%	83.60%	82.00%	77.43%

The SVM and NB models performed adequately but exhibited lower F1-scores, especially when distinguishing between overlapping cognitive classes such as "Cognitively Impaired but not MCI" and "Non-amnestic MCI." DT outperformed others because it effectively handles both categorical and numerical features, which align well with the structure of the Alzheimer's dataset. DT models non-linear relationships and interactions between features and gives an advantage in datasets where cognitive status depends on the complex relationships between demographic, clinical, and test scores.

These results demonstrated two important findings.

1. The training pipeline is stable and capable of producing high-performing models even with reduced data.
2. Model performance is not dependent on the specific sample, suggesting that the features selected are strong predictors of the target class.

Besides its practical contributions, the proposed MindCare application for Alzheimer's disease prediction has several limitations that need further improvement, both in functionality and scope. First, advanced machine learning techniques, such as ensemble methods with Random Forest or Gradient Boosting and deep learning models, further improve predictive performance. Secondly, although the platform provides performance metrics and visualization of selected features, advanced explainability techniques such as SHAP or LIME are not integrated, which may restrict deeper clinical insight into individual predictions. Additionally, data privacy and security must be carefully managed when the application is integrated with the real hospital system. Deployment in real-world clinical environments would require authentication mechanisms and institutional approval. Together, these future enhancements would elevate the application from a proof-of-concept tool into a scalable, production-ready system with meaningful clinical impact.

5 CONCLUSION

MindCare is a web-based AD prediction system that integrates comprehensive data preprocessing pipelines, machine learning algorithms, and a user-friendly interface. The proposed framework utilizes clinical health records, including cognitive scores, demographics, and neuropathological features that are already available in hospital settings. It reduces reliance on specialized infrastructure and costly, invasive imaging scans. This system provides clinicians with data-driven insights to enhance diagnostic accuracy and patient care. The platform not only reports comprehensive performance metrics such as accuracy, precision, recall, and F1 score but also offers visualizations of selected features and classification outcomes. The comparative evaluation of multiple classifiers further strengthens the reliability by not relying on a single model outcome. This approach improves the interpretability of predictive outcomes, facilitating the use of advanced analytics in clinical decision-making. The model can be extended to include multimodal features. The application's modular and scalable design enables it to be adapted for various other medical diagnostic applications.

6 ACKNOWLEDGMENTS

The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADRCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI David Holtzman, MD), P30 AG066518 (PI Lisa Silbert, MD, MCR), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI Julie A. Schneider, MD, MS), P30 AG072978 (PI Ann McKee, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Jessica Langbaum, PhD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Glenn Smith, PhD, ABPP), P30 AG066508 (PI Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P30 AG086401 (PI Erik Roberson, MD, PhD), P30 AG086404 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD).

7 REFERENCES

[1] C. K. K. Reddy *et al.*, "A transfer learning approach: Early prediction of Alzheimer's disease on US healthy aging dataset," *Mathematics*, vol. 12, no. 14, p. 2204, 2024. https://doi.org/10.3390/math12142204

[2] X. Sun, W. Guo, and J. Shen, "Toward attention-based learning to predict the risk of brain degeneration with multimodal medical data," *Frontiers in Neuroscience*, vol. 16, p. 1043626, 2023. https://doi.org/10.3389/fnins.2022.1043626

[3] Alzheimer's Association, "2023 Alzheimer's disease facts and figures," *Alzheimer's Dement.*, vol. 19, no. 4, pp. 1598–1695, 2023. https://doi.org/10.1002/alz.13016

[4] D. B. Hammers *et al.*, "Differences in baseline cognitive performance between participants with early-onset and late-onset Alzheimer's disease: Comparison of LEADS and ADNI," *Alzheimer's & Dementia*, vol. 21, no. 1, p. e14218, 2025. https://doi.org/10.1002/alz.14218

[5] W. H. Abdulsalam, Z. H. Ibrahim, B. H. Majeed, and H. T. Salim ALRikabi, "Utilizing machine learning techniques to predict university students' digital competence," *International Journal of Engineering Pedagogy*, vol. 15, no. 3, pp. 75–91, 2025. https://doi.org/10.3991/ijep.v15i3.54943

[6] S. Akter, Z. Liu, E. J. Simoes, and P. Rao, "Using machine learning and electronic health record (EHR) data for the early prediction of Alzheimer's disease and related dementias," *The Journal of Prevention of Alzheimer's Disease*, vol. 16, p. 100169, 2025. https://doi.org/10.1016/j.tjpad.2025.100169

[7] A. Y. Mahmoud, D. Neagu, D. Scrimieri, and A. R. Abdullatif, "Early diagnosis and personalised treatment focusing on synthetic data modelling: Novel visual learning approach in healthcare," *Computers in Biology and Medicine*, vol. 164, p. 107295, 2023. https://doi.org/10.1016/j.compbiomed.2023.107295

[8] J. A. Pillai, J. Bena, B. Tousi, K. Rothenberg, C. D. Keene, and J. B. Leverenz, "Lewy body pathology modifies risk factors for cerebral amyloid angiopathy when comorbid with Alzheimer's disease pathology," *Alzheimer's & Dementia*, vol. 20, no. 4, pp. 2564–2574, 2024. https://doi.org/10.1002/alz.13704

[9] J. Sung, Z. Jiang, and Y. Huang, "Application and prospects of traditional Chinese medicine in the treatment of Alzheimer's Disease," *IETI Transactions on Data Analysis and Forecasting*, vol. 2, no. 4, pp. 33–45, 2025. https://doi.org/10.3991/itdaf.v2i4.53111

[10] K. A. Sanchez Villogas, P. M. Pinzás Riveros, P. Castañeda, and A. Oñate-Andino, "An artificial intelligence-driven mobile application for real-time assistance in taxi-related criminal incidents using natural language processing in metropolitan lima," *International Journal of Interactive Mobile Technologies*, vol. 20, no. 2, 2026. https://doi.org/10.3991/ijim.v20i02.57541

[11] S. U. Rehman *et al.*, "AI-based tool for early detection of Alzheimer's disease," *Heliyon*, vol. 10, no. 8, p. e29375, 2024. https://doi.org/10.1016/j.heliyon.2024.e29375

[12] Y. Fan, "A web-based tool for early prediction of Alzheimer's disease dementia based on magnetic resonance imaging data," *Alzheimer's & Dementia*, vol. 20, p. e094150, 2024. https://doi.org/10.1002/alz.093505

[13] F. Bertini, D. Beltrami, P. Barakati, L. Calza, E. Ghidoni, and D. Montesi, "A web-based application for screening Alzheimer's disease in the preclinical phase," in *2023 IEEE Symposium on Computers and Communications (ISCC)*, 2023, pp. 1–6. https://doi.org/10.1109/ISCC58397.2023.10218229

[14] G. Balamuralikrishnan, S. Boomika, and R. Sri Nandhini, "A predictive and adaptive mobile/web application for Alzheimers disease using machine learning," in *2024 9th International Conference on Communication and Electronics Systems (ICCES)*, 2024, pp. 1580–1585. https://doi.org/10.1109/ICCES63552.2024.10860053

[15] arpy8. "Alzheimers_prediction_system: Alzheimer's prediction system deployed on Streamlit that employs logistic regression to classify whether a person is prone to having Alzheimer's disease or not [Internet]," *GitHub*. 2023. [Online]. Available: https://github.com/arpy8/Alzheimers_Prediction_System

[16] J. Soriano Reos, "AlzheimInk: A mobile web application for detecting Alzheimer's disease through handwriting analysis," Universitat Oberta de Catalunya (UOC), 2024. [Online]. Available: https://hdl.handle.net/10609/150903

[17] S. Jahan *et al.*, "Explainable AI-based Alzheimer's prediction and management using multimodal data," *PLoS ONE*, vol. 18, no. 11, p. e0294253, 2023. https://doi.org/10.1371/journal.pone.0294253

[18] E. Colantuoni, G. Surplus, A. Hackman, H. M. Arrighi, and R. Brookmeyer, "Web-based application to project the burden of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 6, no. 5, pp. 425–428, 2010. https://doi.org/10.1016/j.jalz.2010.01.014

[19] S. Khan, A. Khan, N. Hameed, M. A. Taufiq, and S. Riaz, "Localizing epileptogenic zone from high density EEG data using machine learning," *Int. J. Online Biomed. Eng.*, vol. 17, no. 6, pp. 73–86, 2021. https://doi.org/10.3991/ijoe.v17i06.18653

[20] A. Khan, L. S. Liu, M. Usman, and S. Fong, "Early diagnosis of Alzheimer disease using instance-based learning techniques," *Journal of Medical Imaging and Health Informatics*, vol. 6, no. 4, pp. 1111–1118, 2016. https://doi.org/10.1166/jmihi.2016.1808

[21] A. Khan and M. Usman, "Alzheimer's disease prediction model using demographics and categorical data," *Int. J. Online Biomed. Eng.*, vol. 15, no. 15, pp. 96–109, 2019. https://doi.org/10.3991/ijoe.v15i15.11472

[22] E. Friedman and L. O. Gostin, "The United Nations sustainable development goals: Achieving the vision of global health with justice," *Georgetown Public Policy Review*, vol. 21, no. 1, 2016.

[23] V. Hassija *et al.*, "Interpreting black-box models: A review on explainable artificial intelligence," *Cognitive Computation*, vol. 16, no. 1, pp. 45–74, 2024. https://doi.org/10.1007/s12559-023-10179-8

[24] L. Besser *et al.*, "Version 3 of the national Alzheimer's coordinating center's uniform data set," *Alzheimer's Disease & Associated Disorders*, vol. 32. no. 4, pp. 351–358, 2018. https://doi.org/10.1097/WAD.0000000000000279

8 AUTHORS

Aunsia Khan is a PhD scholar at the School of Computer Sciences, Universiti Sains Malaysia (USM), Malaysia. Her research interests include data mining, machine learning, decision support systems, image analysis, and multimodal data. She has been the winner of the 3MT competition at her university level and represented her university in a national-level competition (E-mail: aunsia@student.usm.my).

Khausaalyaah Sinathurai is the coordinator of community engagement division at Universiti Sains Malaysia. She has her bachelor's degree in computer sciences and strong interests in machine learning and web/mobile application development related to healthcare and society.

Anusha Achuthan is currently a senior lecturer at the School of Computer Sciences, Universiti Sains Malaysia. Her research interests include the areas of computational neuroscience, knowledge-guided image analysis, biomarker analysis for neurological disorders and aging, machine learning, and computer vision (E-mail: anusha@usm.my).

Anton Lord is an experienced postdoctoral researcher with a PhD. degree in neuroscience from The University of Queensland, Australia. He worked as a principal data scientist and is now the acting manager at Metro South Health. He is skilled in bioinformatics, research design, machine learning, medical research, and spectral analysis.