Health and Technology
https://doi.org/10.1007/s12553-025-01006-1

ORIGINAL PAPER

Cognitive feature evaluation for disease progression in dementia and its precursors using feature selection

Fadi Thabtah[1] · **David Peebles**[2]

Received: 13 June 2024 / Accepted: 9 July 2025

Abstract

Purpose Dementia is a condition with symptoms of memory decline, cognitive impairment, and difficulties in language and problem-solving, among others. Early screening of dementia conditions such as Alzheimer's disease (AD) is fundamental for quick intervention and disease management. Currently used neuropsychological assessments are time-consuming as they contain many elements and require critical resources which are not always available. Other pathological assessments are invasive and not cost effective, hence identifying cognitive features for different dementia sub-groups during progression of the condition is crucial for clinicians. This study investigates this problem using a cost-effective data driven approach.

Methods Using real cases and controls from the Alzheimer's Disease Neuroimaging Initiative data repository (ADNI) who have undergone the Alzheimer's Disease Assessment Scale-Cognitive 13 (ADAS-Cog), we conduct a feature-feature assessment together with Permutation Feature Importance (PFI) and machine learning algorithms to derive influential cognitive features for specific dementia groups from baseline diagnosis up to 36 months.

Results Feature-feature analysis showed correlations between memory tasks such as Word Recall, Delayed Word Recall, and Word Recognition across both CN-MCI and MCI-AD groups. In contrast, low correlations for Naming, Command, and Ideational Praxis suggest they tap into distinct DSM-5 domains thus making them ideal for early screening. In addition, PFI results showed that Delayed Word Recall emerged as a top cognitive marker of progression in early stages, while Orientation gained prominence later thereby reflecting a shift toward executive and attentional decline.

Conclusions The results of this study identified important relationships between cognitive features in the ADAS-Cog and provide a clear example of the value of data-driven machine learning approaches in the identification of markers that indicate disease progression in dementia.

Keywords Dementia · Machine learning · Screening · Cognitive features · Dementia progression

1 Introduction

Dementia is a condition characterised by difficulties in memory, disturbance in language, psychological changes, and sometimes impairments in activities of daily living [1]. Dementia has significant negative impacts on individuals, society and the economy [2] with cases predicted to increase worldwide, for example in the UK to approximately 1.6 million people by 2040 [3]. This current and predicted trend has raised calls for new research initiatives for treatment and diagnosis [4, 5]. One important research area is the improvement of dementia pre-diagnosis using data-driven techniques. The promise of this research is a cost effective, validated, and reliable method of dementia screening and prognosis using large datasets and methods grounded in artificial intelligence and machine learning.

A primary assessment for determining early dementia involves measuring the cognitive abilities that underlie routine activities, e.g., remembering, thinking, problem solving, decision making, and judgement [6]. Some studies have reported that a decline in cognitive skills can indicate early dementia conditions such as Alzheimer's

✉ David Peebles
d.peebles@hud.ac.uk

Fadi Thabtah
f.fayez@adsm.ac.ae

[1] Abu Dhabi School of Management, Abu Dhabi, UAE

[2] Department of Social and Psychological Sciences, University of Huddersfield, Huddersfield HD1 3DH, UK

Published online: 08 August 2025

disease (AD) [7] and two requirements for diagnosing dementia or major neurocognitive disorder are (a) a substantial decline in one of the six cognitive domains defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM−5) framework, and (b) that the cognitive deficits experienced interfere with the performance of everyday activities [8]. The DSM−5 is one of the standard references used by diagnosticians and medical researchers to diagnose and classify mental health conditions, including neurocognitive disorders such as dementia. It offers detailed diagnostic criteria for different types of dementia, including AD, vascular dementia, etc., supporting assessment and treatment planning in clinical settings.

While many studies have investigated dementia diagnosis using cognitive features, identifying which features can be used to measure disease advancement has been less studied, particularly using machine learning [9]. Achieving this involves designing a feature selection process to identify features in a dataset that have high correlations with the target class in an automated manner [10]. Feature selection can simplify data analysis by reducing the input data dimensionality and by pinpointing influential items, in the current context, cognitive items which can be important signals for clinicians during the early clinical evaluation of patients. More importantly, it is also imperative to distinguish between influential cognitive items for the subgroups of dementia, especially pre-dementia conditions like Mild Cognitive Impairment (MCI) and mild dementia, as these are more challenging than the later stages. Unfortunately, most of the existing neuropsychological methods used for dementia pre-diagnosis rarely measure the correlations between cognitive features that would be useful to clinicians [11–13].

In this study we aim to demonstrate the process of identifying key cognitive features, if any, during AD progression, using a dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu/data-samples/adni-data/). The data are records from cases and controls who underwent an Alzheimer's Disease Assessment Scale-Cognitive 13 (ADAS-Cog) assessment [14]. The feature associations were identified by performing feature-feature assessment and PFI using machine learning algorithms on the cognitive features for two dementia groups from baseline diagnosis up to 36 months:

1. Cog-CN-MCI – participants with a baseline diagnosis of Cognitively Normal (CN) who then progressed to MCI.
2. Cog-MCI-AD – participants with a baseline diagnosis of MCI who then progressed to AD.

To accomplish this, we conducted several analyses of the 'ADNI-Merge-ADAS-Cog' dataset in the ADNI repository and established a set of analysis criteria to ascertain potential cognitive features that could be used as signs of the progression of dementia. The research questions that this study aims to answer are:

1. How can the important neuropsychological features, at each dementia stage during progression, be determined using feature selecction analysis?
2. Do cognitive features vary when the dementia stage changes?

In the rest of this paper we contextualise our work in the previous relevant research conducted by ourselves and others and then discuss the methods and data used in our study. Finally, we describe the results of our study and discuss its implications.

2 Related work

Previous studies have sought to identify cognitive features in the ADNI dataset. For example [9], studied cognitive elements in the Clinical Dementia Rating Scale Sum of Boxes (CDR-SB) [15]. The authors developed a data process to assess CDR-SB cognitive features and demographics, and identified the top-ranked features using classification techniques to predict CDR-SB scores. The authors demonstrated that most classification techniques produced models with acceptable predictive power, but that probabilistic and Decision Tree (DT) (C4.5) algorithms were more appropriate to predict CDR-SB scores using the considered baseline data.

A number of studies have used machine learning to determine which functional activities are related to dementia progression and indicate AD. For example, using feature selection with classification algorithms on data from the Functional Activity Questionnaire (FAQ) [16] in the ADNI database [17], Thabtah et al., [18], derived highly predictive models and revealed that functional activities such as 'Shopping' and 'Administration' showed good association with dementia and possible progression. Expanding on this functional assessment [19], sought to identify influential cognitive elements of the ADAS-Cog using a data-driven approach consisting of feature selection and classification. The authors used three feature selection methods, including Information Gain, ReliefF, and Chi-square testing along with various classification algorithms including DTs (C4.5), Bayesian Networks, and Logistic Regression on data from the ADNI database. The results showed that processing only

three cognitive elements by the classification algorithms, particularly C4.5, revealed high predictive classification models. Overall, the results of [18, 19] showed that cognitive elements are more indicative of early dementia progression than functional activities. However, these studies did not consider dementia sub-groups during the process of learning or feature selection or how features could change during the progression stages of the disease.

Other studies have evaluated previously designed dementia tests to determine the most predictive cognitive and functional factors. For example [6], investigated quality issues (clinical relevance, validity) in the Cognitive-Functional Composite (CFC), a composite battery test for mild dementia comprising multiple cognitive tests and one functional ability questionnaire [20]. Analysing a dataset of 184 subjects with different dementias and pre-dementia conditions (MCI, AD, Lewy body), the authors showed that the CFC's main components (executive functions, episodic memory, and Amsterdam Instrumental Activities of Daily Living (A-IADL) score) were more highly correlated with disease severity than conventional cognitive tests such as CDR-SB score and ADAS-Cog13 score.

Several researchers have compared different versions of the ADAS-Cog. For example [21], conducted a review which found the ADAS-Cog13 [22] to be better at identifying disease progression in patients with AD than ADAS-Cog11. In another study [23], used ADNI data to compare the performance of the 3-, 5-, 11-, and 13-item ADAS-Cog variants on their ability to detect cognitive decline. While the original 11-item ADAS-Cog was able to measure cognition in patients with mild to moderate AD, it was unable to detect change or measure cognitive domains known to cause impairment in early-stage AD. The additional tests such as digit cancellation and delayed word recall have given the 13-item ADAS-Cog improved ability for early AD screening. Overall, the authors concluded that the impact of expansion or reduction of the ADAS-Cog was subtle, but noted that in mild AD, adding rather than removing items was more beneficial.

More recent studies have further evaluated the reliability and interpretability of ADAS-Cog13 [24]. conducted a systematic review of minimal clinically important differences (MCIDs) for ADAS-Cog, concluding that changes of + 2 to + 3 points in MCI and + 3 points in mild AD should be interpreted as clinically meaningful. These thresholds provide crucial benchmarks for interpreting progression and evaluating treatment effects in clinical trials [25]. used a latent state–trait model with autoregressive effects to assess whether ADAS-Cog13 items reflected state-specific, trait-like, or accumulated features over time. They found that while some language and memory items (e.g., naming and word recall) had high reliability and reflected gradual accumulation of cognitive dysfunction, most items showed poor psychometric reliability, complicating the interpretation of change scores in longitudinal analyses. In addition [26], applied Bayesian latent class modelling to assess the diagnostic performance of ADAS-Cog13, MMSE, and MoCA. ADAS-Cog13 outperformed the other tests in both sensitivity and specificity when detecting both MCI and AD, confirming its continued relevance for clinical screening and research.

[27] used ADAS-Cog13 and other clinical variables in a machine learning process to predict 2-year cognitive decline in early AD patients. Their model achieved significant accuracy, demonstrating the value of baseline cognitive data for forecasting progression trajectories and their potential benefit for clinical trial enrichment and the monitoring of individual patient trajectories [28]. developed a multimodal multitask deep learning model trained on ADNI data to predict ADAS-Cog13 and MMSE scores along with diagnostic status. Their model achieved higher accuracy and lower error than previous approaches, reinforcing the usefulness of ADAS-Cog13 in multimodal machine learning models.

Finally [29], assessed whether cognitive measures could be used to reduce the number of neuropsychological tests required. Using both computational and human expert feature reduction approaches, they found that three ADAS-Cog13 items were sufficient to build high-accuracy classifiers using support vector machines. Their results reinforce the potential of ADAS-Cog13 not just as a total score but as a structured set of features with differing diagnostic values. Also worth noting is a study by [30], which mapped cognitive dysfunction across ADAS-Cog, CDR-SB, and MMSE [14, 15, 31] using item response theory (IRT) [32]. The results indicated that ADAS-Cog and CDR-SB yielded consistent estimates of dysfunction across various stages, further validating their use in parallel for tracking progression. Together, these studies demonstrate that although ADAS-Cog13 has some limitations, particularly in its aggregate scoring, it remains a crucial tool for detecting cognitive decline and tracking Alzheimer's disease progression, especially when its individual items and psychometric properties are properly modelled.

3 Methods, data & experimental platform

3.1 Methods

This study used the ADAS-Cog−13, a standard clinical tool for assessing cognitive dysfunction [14], comprising 13 tasks spanning memory, language, praxis, and orientation. It is scored from 0 to 85, with higher scores indicating greater impairment. While its sensitivity in early dementia is debated [20], it remains widely used. Prior work identified a cut-off score ≥ 12 as optimal for AD detection, with 89.2% sensitivity and 88.5% specificity [33]. To evaluate cognitive item relationships, we performed a feature-to-feature correlation analysis using Pearson's r, excluding the target class. This approach has been successfully applied in previous dementia-related studies [e.g., 34, 35, 18, 19] and highlights redundant features by identifying items with high intercorrelations (r near ± 1), allowing us to focus on distinct contributors to dementia progression. Results were mapped to DSM-5 cognitive domains to ensure broad domain coverage during feature selection.

We conducted a feature-to-feature assessment (excluding the target class) to explore inter-item relationships among ADAS-Cog-13 cognitive tasks. This analysis identifies highly correlated items that may convey redundant information regarding disease progression, allowing for dimensionality reduction. When two features exhibit strong correlation, their contribution to classification is likely similar, and one may be omitted without loss of predictive value. Cognitive items were also mapped to DSM-5 domains to ensure broad diagnostic coverage during feature selection.

We computed Pearson correlation coefficients between features, summarised in matrix form. This highlights similarities across tasks and supports identification of distinctive, diagnostically relevant items. The correlation coefficient *r*, which ranges from − 1 to + 1, was calculated as:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}} \quad (1)$$

.

where n is the number of observations, x and y are feature values, and $\rightarrow x$ and $\rightarrow y$ are their respective means. Higher absolute values of r indicate stronger linear dependence between features.

We also applied Permutation Feature Importance (PFI) [36] to quantify the influence of each ADAS-Cog item. PFI estimates the drop in model accuracy when a feature is randomly shuffled, revealing its predictive value. We implemented PFI using two classifiers: DT (C4.5) and LR, both chosen for their differing learning approaches and interpretability in medical research. To complement PFI, we used Mutual Information (MI) [37], a measure of non-linear dependence between each cognitive feature and the class label. MI is calculated as:

$$MI(X,Y) = \sum\nolimits_{x,y} p(x,y) \cdot log\left(\frac{p(x,y)}{p(x) \cdot p(y)}\right) \quad (2)$$

where $p(x, y)$ is the joint distribution of features X and Y. Higher MI values indicate stronger relevance for classification. All analyses were implemented in Python via Google Colab using the Scikit-learn library for model construction, PFI, and MI experiments. This computational setup ensured reproducibility and scalability across multiple dementia progression cohorts.

3.2 Data and Preparation

This study used data from the ADNI data repository [38], specifically the ADNI-Merge and ADAS-Cog-13 datasets. ADNI-Merge aggregates data across several ADNI phases (ADNI-1, GO, 2, and 3), and includes a wide range of participant data such as demographics, clinical diagnoses, neuropsychological assessments, functional questionnaires, genetic markers (e.g., APOE status), and biomarker data from structural MRI and PET imaging. The ADAS-Cog-13 dataset contains scores for 13 individual cognitive tasks measuring memory, language, praxis, and orientation. Table 1 displays the basic statistics related to the ADNI-Merge, and ADAS-Cog datasets. There is one target class in the ADNI-Merge dataset, which is the diagnosis of the last examination visit (DX); another attribute that can be considered important is the baseline diagnosis (DX_bl), which denotes the initial diagnosis given to the patient at the first visit.

Table 1 General statistics of the datasets before Pre-processing

Dataset Name	# of Features	# of Patients	# of Data Observations (visits)	Missing Values in Key Attributes
ADNIMERGE	113	2,260	14,627	Class DX: 4,243 missing values
ADAS-Cog	121	1,751	6,770	100 missing values: 91 across ADAS tasks and 9 in VISCODE2 attribute

To investigate cognitive changes during the progression of dementia, we created two longitudinal sub-cohorts by merging the ADAS-Cog-13 scores with diagnostic data from ADNI-Merge:

- The **CN to MCI cohort** consists of participants who were diagnosed as cognitively normal (CN) at their baseline visit but later progressed to mild cognitive impairment (MCI) within a 36-month follow-up period.
- The **MCI to AD cohort** includes participants who were diagnosed with MCI at baseline and subsequently progressed to Alzheimer's disease (AD) over the same timeframe.

To construct these sub-cohorts, we first unified the ADAS Sub-Scores and Total Scores datasets from the various ADNI phases to ensure full coverage of ADAS-Cog-13 items. We then left-joined this cognitive data with the ADNI-Merge dataset based on unique participant IDs (RID) and visit codes (VISCODE). Visits were excluded if either the diagnosis or ADAS-Cog scores were missing. Importantly, we only retained participants who demonstrated diagnostic progression (from CN to MCI or from MCI to AD) and excluded all visits that occurred after the progression point to avoid conflating pre- and post-progression cognitive data.

After these exclusions, the resulting CN to MCI dataset consisted of 141 participants contributing 744 observations (visits), while the MCI to AD dataset comprised 385 participants and 1,765 observations. However, in both datasets the majority of observations represented no change in diagnosis. For example, only 73 of the 1,695 visits in the CN to MCI dataset (about 4.3%) indicated progression, while 207 of 3,275 visits in the MCI to AD dataset (6.3%) reflected diagnostic change. This large class imbalance posed a risk of biased model training and reduced classification performance. To address this, we applied the Synthetic Minority Over-Sampling Technique (SMOTE) [39], using the UBL package for R [40] with $k = 5$ nearest neighbours. SMOTE generates synthetic examples of minority class instances (i.e., those showing progression) based on feature space similarity, resulting in more balanced datasets suitable for machine learning analysis. After resampling, the CN to MCI cohort included 3,198 observations, with approximately 49.3% progression cases, while the MCI to AD cohort included 5,982 observations, with 48.7% progression cases. Table 2 displays the statistics of the participant groups for 36 months from the baseline visit and for different dementia stages.

During dataset preparation, we identified a small number of regression cases, in which participants appeared to revert to a less severe diagnosis over time (e.g., MCI to CN or AD to MCI). Specifically, 20 regression cases were found in the CN to MCI cohort and 16 in the MCI to AD cohort (labelled '-1'). These were excluded from analysis because they represent atypical patterns that fall outside the scope of this study, which focuses exclusively on forward disease progression. These balanced cohorts provided the basis for the feature correlation, importance ranking, and classification analyses reported in the following sections.

3.3 Analysis method

The analyses were conducted using Python [41] using hyperparameters of the feature selection method that were unchanged. We assessed the feature-to-feature correlation within the datasets and identified highly correlated items in order to derive influential features from the 'ADNI-Merge-ADAS-Cog' dataset. Feature-feature correlation matrices with coefficients indicating the strength between two items were created to identify highly correlated features by calculating the largest mean absolute correlation between each item to remove any redundant features. Pearson correlation was used to generate a correlation matrix of the data's features as a vector of integers to reduce independent attributes' correlations. When two attributes are highly correlated, the function

Table 2 General statistics for the groups of participants within 36 months from the baseline and for different dementia stages

Dataset Name	# of Patients before sampling	# of Data Observations (visits)	DX Progress - Class Distribution before Data Balancing	DX Progress - Class Distribution after Data Balancing
ADNI-Merge-ADAS-Cog CN to MCI	287	1,695	Total observations: 1,695 '0': 1,622 (majority 95.70%) '1': 73 (4.30%) −1:20	Total observations: 3,198 '0': 1,622 (50.71%) '1': 1,576 (49.28%)
ADNI-Merge-ADAS-Cog MCI to AD	651	3,275	Total observations: 3,275 '0': 3,068 (majority 93.68%) '1': 207 (6.32%) −1:16	Total observations: 5,982 '0': 3,068 (51.29%) '1': 2,914 (48.71%)

evaluates the correlation of the mean absolute value for each attribute and drops the one with the greatest value. The suggested Cut-off = 0.60 [42]. For data balancing the SMOTE algorithm was used to sample the minority class labels in the dataset. SMOTE is a data sampling technique that adjusts the class distribution by taking the entire dataset as input, thereby increasing the minority class using K Nearest Neighbours (KNN) [43].

4 Results and analysis

4.1 Correlation analysis

We performed feature ranking using feature selection methods alongside feature-feature assessment on the cognitive items for the CN-MCI and MCI-AD groups from baseline diagnosis up to 36 months. Feature-to-feature correlation matrices (Figs. 1 and 2) revealed strong associations between Word Recall and Delayed Word Recall in both cohorts ($r = 0.69$ and 0.59, respectively), as well as moderate correlations with Word Recognition. These overlaps suggest shared variance among memory tasks. For example, correlations between Delayed Word Recall and Word Recognition were 0.41 (CN–MCI) and 0.51 (MCI–AD), demonstrating consistent interdependence during progression. In contrast, Command, Naming, Ideational Praxis, and Word-finding Difficulty showed weak or negative correlations with other items in the CN–MCI group, suggesting that they assess distinct cognitive domains. These features map onto DSM-5 domains: language, perceptual motor function, learning

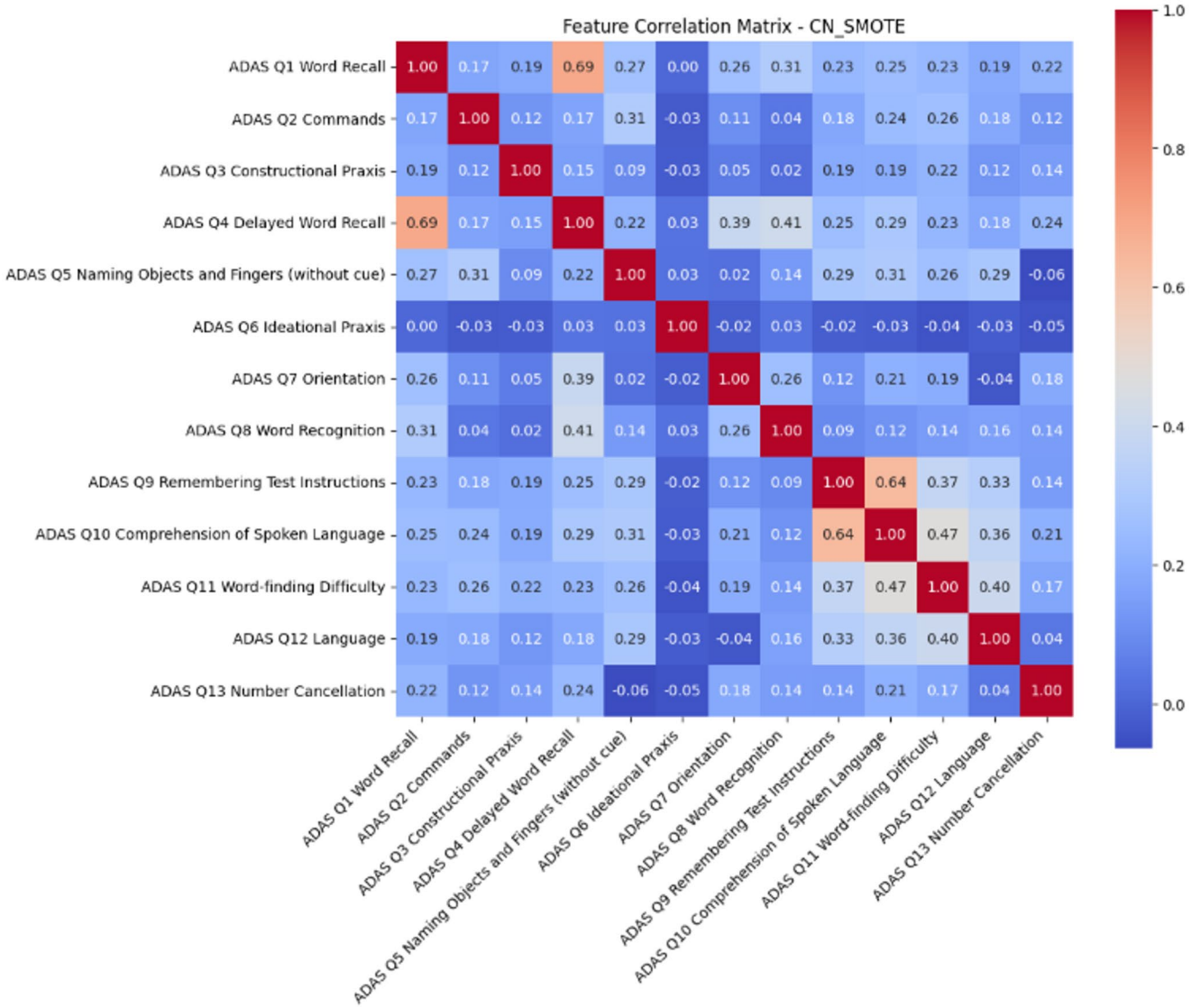

Fig. 1 Feature-Feature Correlation Matrix of Cognitive Items for CN to MCI Group

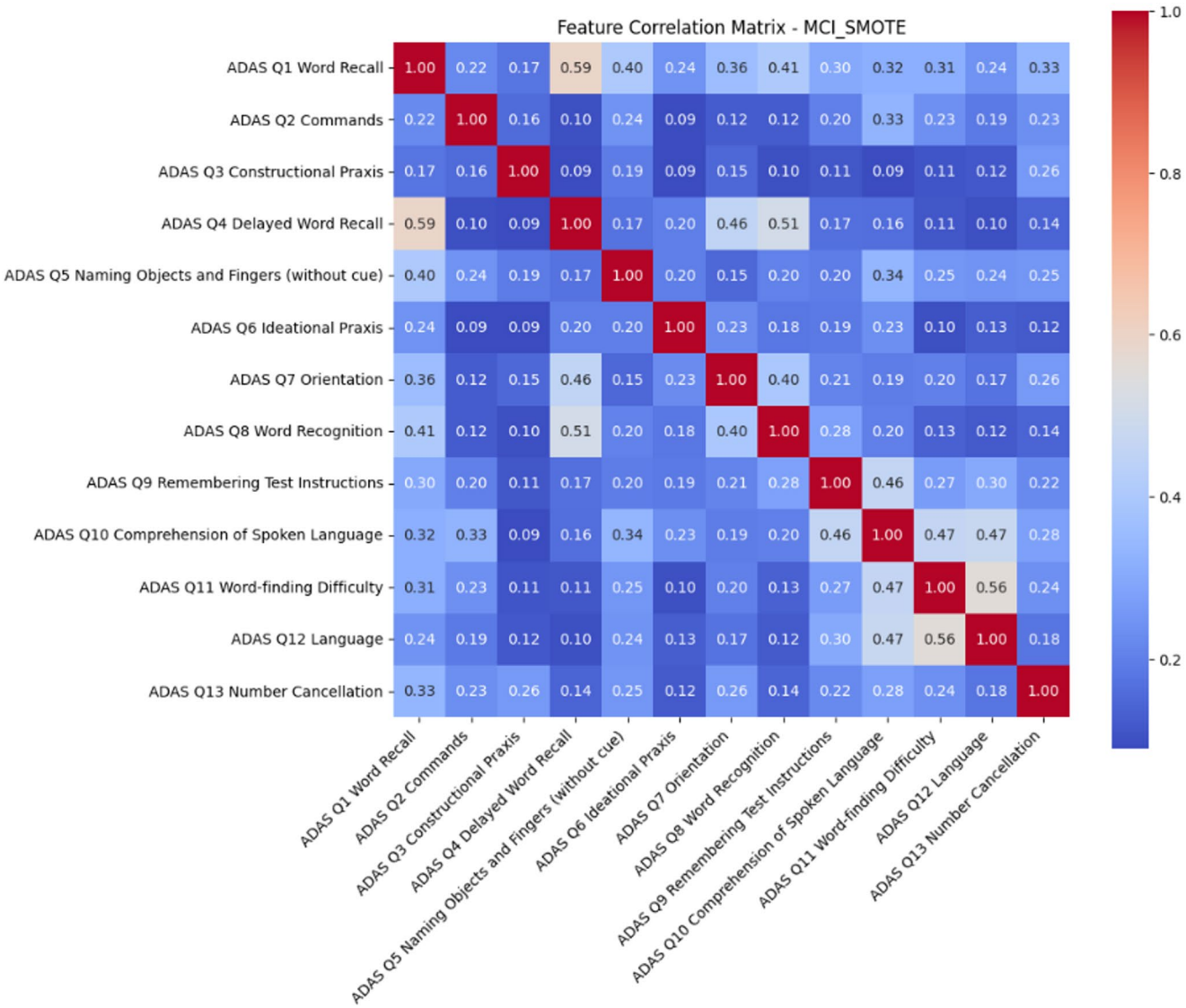

Fig. 2 Feature-Feature Correlation Matrix of Cognitive Items for MCI to AD Group

and memory, and executive function/complex attention, making them potentially valuable for early-stage screening. In the MCI–AD group, inter-feature correlations increased among memory and comprehension tasks, thus reflecting more global cognitive decline.

4.2 Permutation feature importance (PFI)

Figures 3A and 4A display PFI rankings using DT and LR classifiers for the CN-MCI cohort. Delayed Word Recall consistently ranked highest, thus aligning with its role in learning and memory. Word Recognition and Word Recall also appeared as top features, at least for DT models predicting CN-to-MCI progression. Figures 3B and 4B show PFI results for the MCI-AD cohort in which Orientation, Word Recognition, and Delayed Word Recall were identified as key cognitive markers. The increased importance of Orientation (temporal/spatial awareness) suggests a shift toward complex attention and executive function deficits alongside ongoing memory decline. Delayed Word Recall remains critical across both stages but is more dominant in early transitions, whereas Orientation is more significant in later progression. Further, Number Cancellation (linked to Attention) and Naming Objects (linked to Language) appear in both stages in DT models with differing rankings representing progressive cognitive decline across the AD trajectory.

4.3 Mutual information (MI)

Figures 5 and 6 display MI rankings. For CN-MCI (Fig. 5), memory tasks (Word Recall, Delayed Word Recall, Word Recognition) dominated, reinforcing their

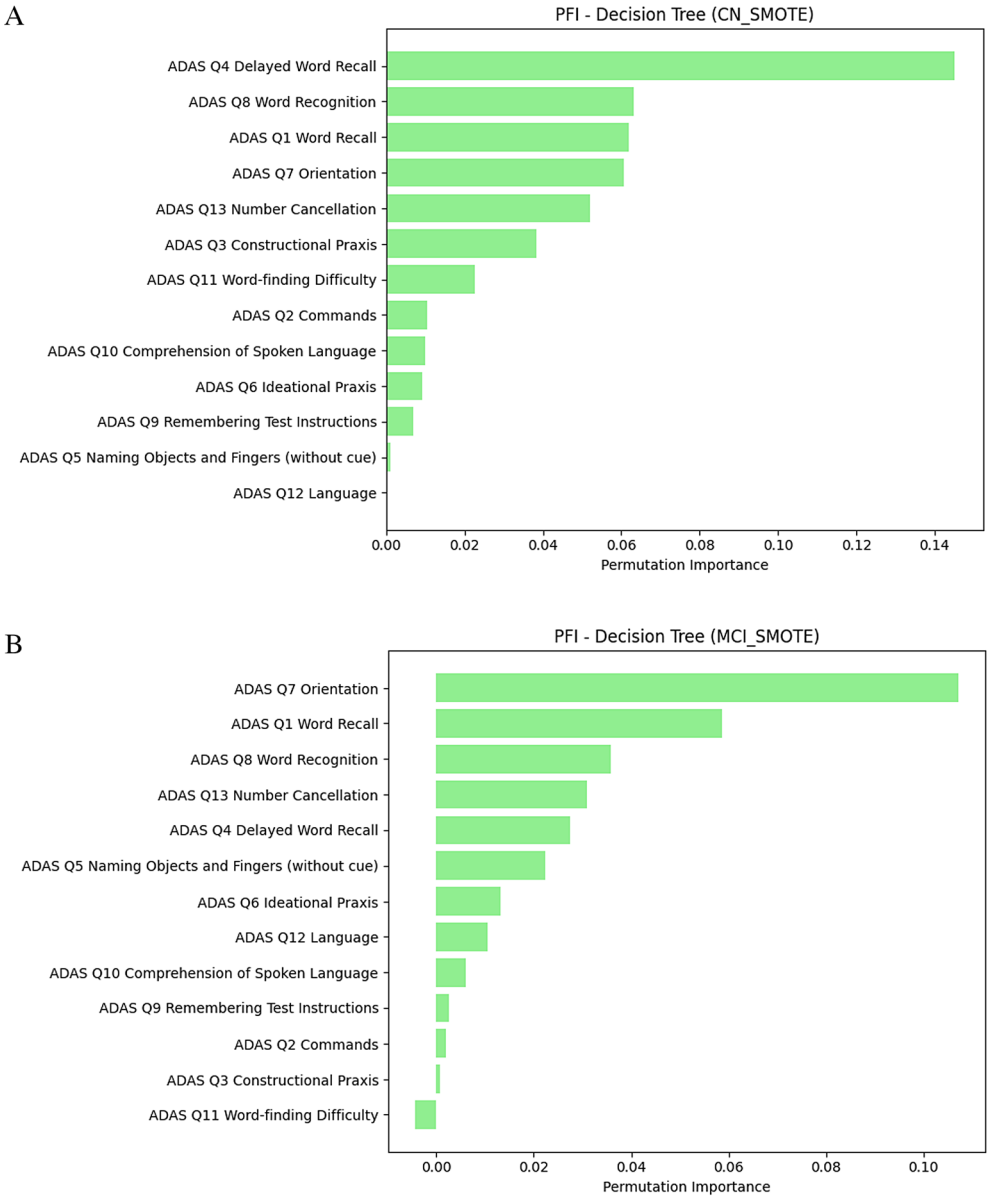

Fig. 3 (**A**) Feature Importance in Permutation % using PFI with a DT classifier on CN-MCI Cohort. (**B**) Feature Importance in Permutation % using PFI with a DT classifier on MCI-AD Cohort

early diagnostic value, while language and praxis items contributed less. These results reflect episodic memory deficits that typically occur during the transition from CN to MCI. In the MCI-AD group (Fig. 6), Word Recognition again ranked highest, but Commands and Comprehension became more predictive, consistent with deterioration in executive and attentional functions. Across both methods and cohorts, Word Recognition consistently emerged as a strong predictor of progression. The differing profiles, especially the reduced relevance of Delayed Word Recall and increased importance of Commands in later stages, highlight evolving cognitive signatures across dementia progression.

5 Conclusions

Identifying cognitive markers of dementia progression, especially in early and intermediate stages, is crucial for timely diagnosis. This study evaluate cognitive features from the ADAS-Cog-13 by focusing on their relevance to disease progression from CN to MCI and from MCI to AD cohorts.

Using correlation analysis and two feature selection methods: PFI and MI, we identified distinct patterns of cognitive decline. In the CN-MCI cohort, memory-related features such as Word Recall, Delayed Word Recall, and Word Recognition emerged as the most informative, thus

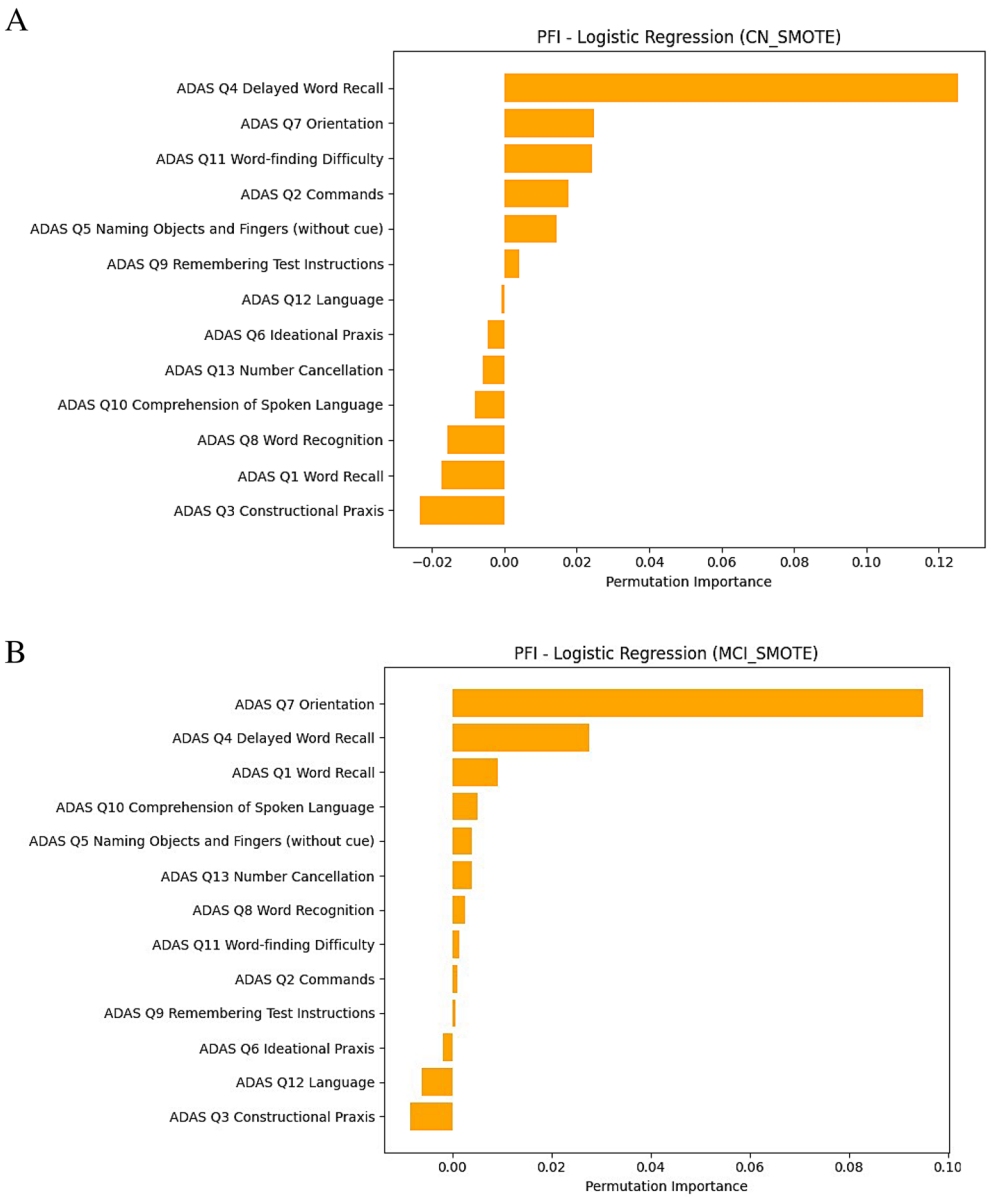

Fig. 4 (**A**) Feature Importance in Permutation % using PFI with a LR Classifier on CN-MCI Cohort. (**B**) Feature Importance in Permutation % using PFI with a LR Classifier on MCI-AD Cohort

aligning with early-stage deficits in learning and memory as defined by DSM-5. The MCI-AD cohort showed a broader profile with increased relevance of Command and Comprehension, indicating deterioration across multiple cognitive domains, including complex attention, executive function and language.

A consistent finding across both groups was the central role of Word Recognition, which maintained high predictive value throughout disease progression. Meanwhile, the declining importance of Delayed Word Recall and the rising importance of Orientation highlight a cognitive shift from specific memory impairment to multi-domain decline as dementia advances. PFI and MI rankings further confirmed the dominance of memory tasks during early progression while attention and comprehension features became more relevant in later stages. These evolving patterns highlight the diagnostic value of tracking item-level changes to detect preclinical AD transitions.

While these findings support the use of targeted cognitive assessments for tracking disease progression, the study has limitations. We relied solely on cognitive data from ADAS-Cog-13 and did not incorporate neuroimaging, biomarker, or genetic information. Moreover, the data were limited to the ADNI sample, which may not fully represent the diversity found in clinical populations. Finally, although we used longitudinal data, we

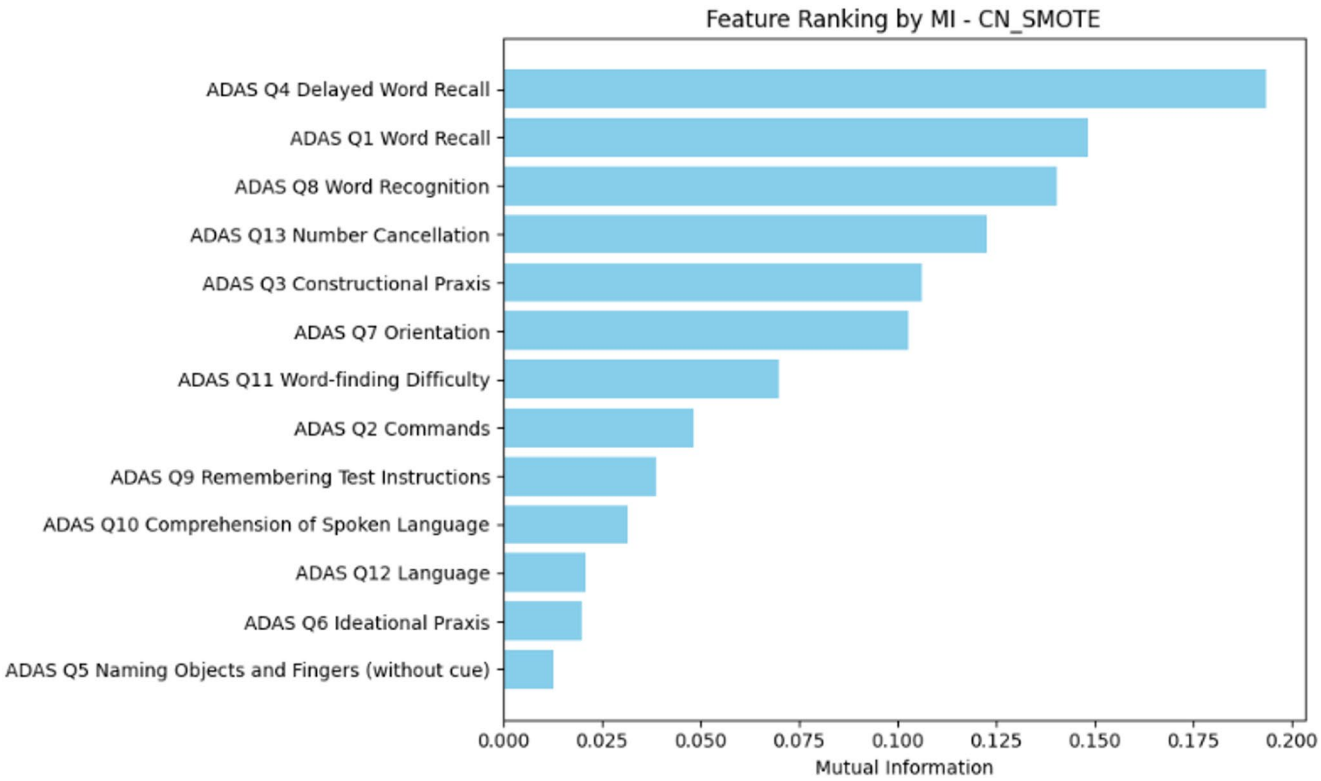

Fig. 5 Feature Ranking Using MI feature selection method on CN-MCI data cohort

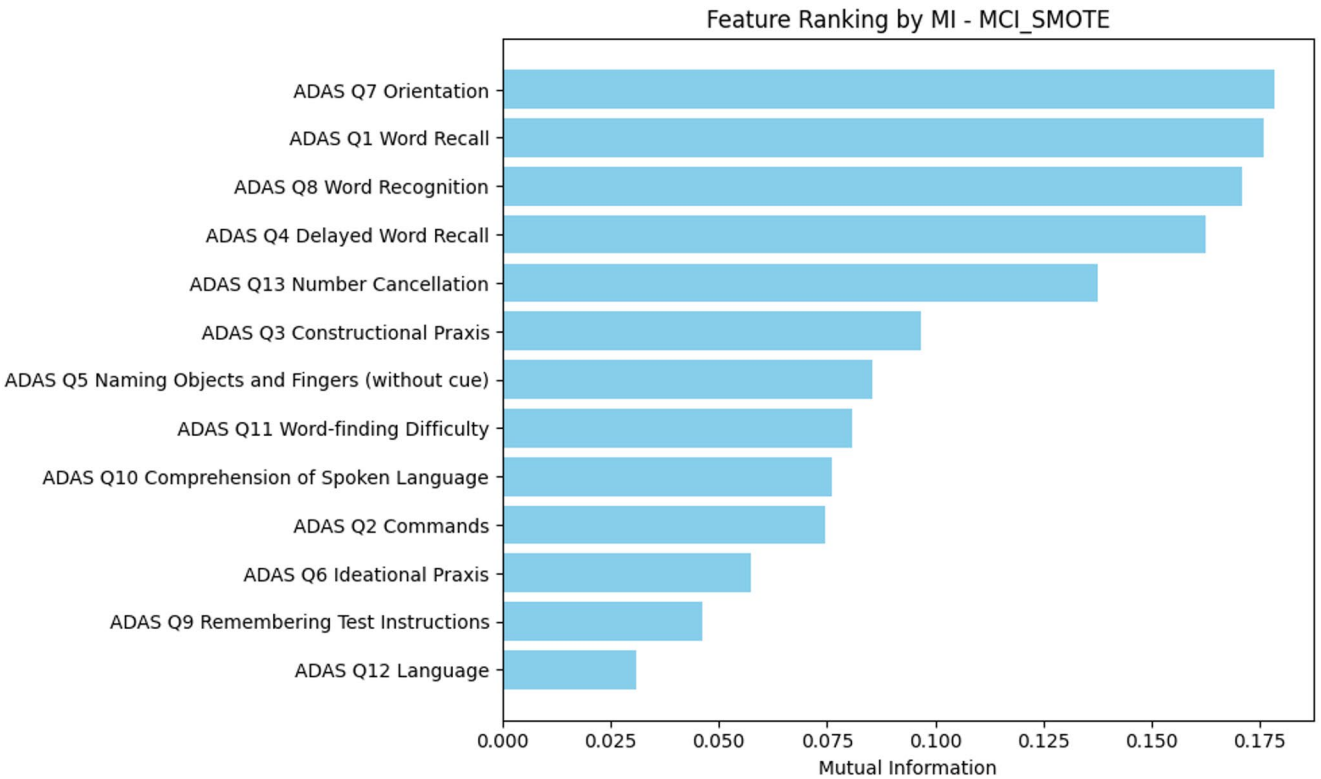

Fig. 6 Feature Ranking Using MI feature selection method on MCI-AD data cohort

did not explicitly model time-based changes, which may be important in capturing trajectories of decline. Future work should incorporate more data sources and consider broader cohorts to enhance generalizability. Incorporating temporal modelling techniques could also help identify more nuanced patterns of progression and improve clinical decision-making.

Acknowledgements Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

Author contributions Conceptualization, Thabtah & Peebles; Methodology, Thabtah; Formal analysis, Thabtah; Investigation, Thabtah; Writing – original draft, Thabtah & Peebles; Writing – review & editing, Thabtah & Peebles; Supervision, Peebles.

Funding This research received no external funding.

Data availability Data are available at: https://adni.loni.usc.edu/.

Code Availability No specialist code was used in this study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This study does not involve humans or animals and so ethical approval is not required.

References

1. Kim H-J, Min J-Y, Min K-B. The association between longest-held lifetime occupation and late-life cognitive impairment: Korean longitudinal study of ageing (2006–2016). Int J Environ Res Public Health. 2020;17(17):6270. https://doi.org/10.3390/ijerph17176270
2. World Health Organization. (2020, September). Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia
3. Wittenberg R, Hu B, Barraza-Araiza L, Rehill A. Projections of older people with dementia and costs of dementia care in the United Kingdom, 2019–2040. The London School of Economics and Political Science, Care Policy and Evaluation Centre. 2019. https://www.alzheimers.org.uk/sites/default/files/2019-11/cpec_report_november_2019.pdf
4. Thabtah F, Spencer R, Peebles D. Common dementia screening procedures: DSM–5 fulfilment and mapping to cognitive domains. International Journal of Behavioural and Healthcare Research. 2022;8(1–2):2022.
5. Pickett J, Bird C, Ballard C, Banerjee S, Brayne C, Cowan K, Clare L, Comas-Herrera A, Corner L, Daley S, Knapp M. A roadmap to advance dementia research in prevention, diagnosis, intervention, and care by 2025. Int J Geriatr Psychiatry. 2018;33(7):900–6.
6. Jutten RJ, Harrison JE, Kjoe LM, Ingala S, Vreeswijk R, van Deelen RAJ, de Jong J, Opmeer F, Aleman EM, Ritchie A, Scheltens CW, P., Sikkes SAM. Assessing cognition and daily function in early dementia using the cognitive-functional composite: findings from the Catch-Cog study cohort. Alzheimer's Res Therapy. 2019;11., Article 45. https://doi.org/10.1186/s13195-019-0500-5
7. Clemmensen FK, Hoffmann K, Siersma V, Sobol N, Andersen BB, Vogel A, Lolk A, Gottrup H, Høgh P, Waldemar G, Hasselbalch SG, Frederiksen KS. The role of physical and cognitive function in performance of activities of daily living in patients with mild-to-moderate alzheimer's disease – a cross-sectional study. BMC Geriatr. 2020;20. https://doi.org/10.1186/s12877-020-01926-9. Article 513.
8. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM–5. 2013. https://www.psychiatry.org/psychiatrists/practice/dsm
9. AlShboul R, Thabtah F, Walter Scott AJ, Wang Y. (2023) The Application of Intelligent Data Models for Dementia Classification. Applied Sciences. 2023;13(6):3612. https://doi.org/10.3390/app13063612
10. Lemoine B, Rayburn S, Benton R. Data fusion and feature selection for Alzheimer's diagnosis. International Conference on Brain Informatics. 2010;6334:320–327. Toronto, ON, Canada: Springer-Verlag. https://doi.org/10.1007/978-3-642-15314-3_30
11. Thabtah F, Spencer R, Ye Y. The correlation of everyday cognition test scores and the progression of alzheimer's disease: a data analytics study. Health Inf Sci Syst. 2020;8:24. https://doi.org/10.1007/s13755-020-00114-8
12. Lim KB, Kim J, Lee HJ, Yoo J, You EC, Kang J. Correlation between Montreal cognitive assessment and functional outcome in subacute stroke patients with cognitive dysfunction. Annals Rehabilitation Med. 2018;42(1):26–34. https://doi.org/10.5535/arm.2018.42.1.26
13. Martyr A, Clare L. Executive function and activities of daily living in alzheimer's disease: A correlational meta-analysis. Dement Geriatr Cogn Disord. 2012;33(2–3):189–203. https://doi.org/10.1159/000338233
14. Rosen W, Mohs R, Davis K. A new rating scale for alzheimer's disease. Am J Psychiatry. 1984;141(11):1356–64. https://doi.org/10.1176/ajp.141.11.1356
15. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry. 1982;160(6):566–72. https://doi.org/10.1192/bjp.140.6.566
16. Pfeffer RI, Kurosaki TT, Harrah CH Jr., Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol. 1982;37(3):323–9. https://doi.org/10.1093/geronj/37.3.323
17. Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust,W.,… Beckett, L. The Alzheimer's disease neuroimaging initiative. Neuroimaging Clinics of North America. 2005;15(4):869–877:xi–xii.
18. Thabtah F, Ong S, Peebles D. Detection of dementia progression from functional activities data using machine learning techniques. Intell Decis Technol. 2022b;16(3):615–30.

19. Thabtah F, Ong S, Peebles D. (2022) Examining cognitive factors for alzheimer's disease progression using computational intelligence. Healthcare 2022, 10, 2045. https://doi.org/10.3390/healthcare10102045
20. Jutten RJ, Harrison JE, de Jong FJ, Aleman A, Ritchie CW, Scheltens P, Sikkes SAM. A composite measure of cognitive and functional progression in alzheimer's disease: design of the capturing changes in cognition study. Alzheimer's Dement. 2017;3(1):130–8. https://doi.org/10.1016/j.trci.2017.01.004
21. Kueper J, Speechley M, Montero-Odasso M. The alzheimer's disease assessment Scale–Cognitive subscale (ADAS-Cog): modifications and responsiveness in pre-dementia populations. A narrative review. J Alzheimer's Disease. 2018;63(2):423–44. https://doi.org/10.3233/JAD-170991
22. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas L, Geldmacher D, Clark C, Thal LJ. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the alzheimer's disease assessment scale that broaden its scope. The alzheimer's disease cooperative society. Alzheimer Dis Assoc Disord. 1997;11(2):S13–21.
23. Podhorna J, Krahnke T, Shear ME, Harrison J, for the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease assessment Scale – Cognitive subscale variants in mild cognitive impairment and mild alzheimer's disease: change over time and the effect of enrichment strategies. Alzheimer's Res Therapy. 2016;8(8). https://doi.org/10.1186/s13195-016-0170-5
24. Muir C, et al. Minimal clinically important difference in alzheimer's disease: rapid review of the literature and implications for clinical trial design. Alzheimer's Dement. 2024;20(1):123–35. https://doi.org/10.1002/alz.13567
25. Cogo-Moreira H, et al. State, trait, and accumulated features of the alzheimer's disease assessment Scale-Cognitive subscale in mild alzheimer's disease: A dynamic structural equation modeling approach. Alzheimer's Dementia: Translational Res Clin Interventions. 2023;9(1):e12453. https://doi.org/10.1002/trc2.12453
26. Wang X, et al. Bayesian Estimation for the accuracy of three neuropsychological tests in detecting alzheimer's disease and mild cognitive impairment: A retrospective analysis of the ADNI database. Eur J Med Res. 2023;28:427. https://doi.org/10.1186/s40001-023-01265-6
27. Devanarayan V, et al. Predicting clinical progression trajectories of early alzheimer's disease: A latent trajectory analysis of ADNI data. Alzheimer's Dement. 2023;19(2):389–403. https://doi.org/10.1002/alz.12846
28. Seshadri N, Shah R, McCalla S. (2021). Early prediction of alzheimer's disease with a multimodal multitask deep learning model. J Student Res, 10(1). https://jsr.org/index.php/path/article/view/1471
29. Battista P, Salvatore C, Castiglioni I. Optimizing neuropsychological assessments for cognitive, behavioral, and functional impairment classification: A machine learning study. Behavioural Neurology, 2017, Article 1850909. 2017.https://doi.org/10.1155/2017/1850909
30. Balsis S, Benge JF, Lowe DA, Geraci L, Doody RS. How do scores on the ADAS-Cog, MMSE, and CDR-SOB correspond? Clin Neuropsychol. 2015;29(7):1002–9. https://doi.org/10.1080/13854046.2015.1119312
31. Folstein M, Folstein SE, McHugh P. Ini-mental state. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189–98. https://doi.org/10.1016/0022-3956(75)90026-6
32. Wright BD. IRT in the 1990s: which models work best? Rasch Meas Trans. 1992;6(1):196–200.
33. Monllau A, Pena-Casanova J, Blesa R, Aguilar M, Bohm P, Sol JM, Hernandez G. Diagnostic value and functional correlations of the ADAS-Cog scale in alzheimer's disease: data on NORMACODEM project. Neurologia. 2007;22(8):493–501.
34. Pereira T, Ferreira F, Cardoso S, Silva D, de Mendonca A, Guerreiro M, Madeira S. Neuropsychological predictors of conversion from mild cognitive impairment to alzheimer's disease: A feature selection ensemble combining stability and predictability. BMC Med Inf Decis Mak. 2018;18(2018):Article137. https://doi.org/10.1186/s12911-018-0710-y
35. Zhu F, Li X, Haipeng T, He Z, Zhang C, Hung G-U, Chiu P-Y, Zhou W. Machine learning for the preliminary diagnosis of dementia. Sci Program. 2020. https://doi.org/10.1155/2020/5629090
36. Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. J Mach Learn Res. 2019;20(177):1–81. http://jmlr.org/papers/v20/18-760.html
37. MacKay DJC. Information theory, inference, and learning algorithms. Cambridge University Press; 2003.
38. Alzheimer's Disease Neuroimaging Initiative [ADNI]. (2021). http://adni.loni.usc.edu
39. Chawla N, Bowyer K, Hall L, Kegelmeyer P. SMOTE: Synthetic Minority Over-sampling Technique. In International Conference of Knowledge Based Computer Systems. 2000:46–57. National Center for Software Technology, Mumbai, India, Allied Press.
40. Branco P, Ribeiro RP, Torgo L. UBL: an R package for utility-based learning, arXiv Prepr. arXiv1604.08079 (2016) [Online]. Available: http://arxiv.org/abs/1604.08079
41. Van Rossum G, Drake FL Jr. Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam; 1995.
42. Akoglu H. User's guide to correlation coefficients. Turkish J Emerg Med. 2018;18(13):91–3. https://doi.org/10.1016/j.tjem.2018.08.001
43. Aha D, Kibler D, Albert MK. Instance-based learning algorithms. Mach Learn. 1991;6:376–6. https://doi.org/10.1007/BF00153759