

Xana Cid-Mejías^{1,2,3*} MSc, Sonia Valladares-Rodríguez⁴ PhD, Alba Felpete^{1,2} PhD, Fátima Fernández-Feijoo^{1,2,3} MSc, Ingri K. Sandoval-García^{1,2,3} MSc, Jose M. Aldrey MD⁵, Marta Saavedra-Piñeiro⁶ PhD, Arturo X. Pereiro^{1,2,3} PhD.

¹Instituto de Psicología (IPsiUS), Universidade de Santiago de Compostela, Santiago de Compostela, A Coruña, Spain

²Applied Cognitive Neuroscience and Psychogerontology group, Health Research Institute of Santiago (IDIS), Santiago de Compostela, A Coruña, Spain

³Department of Developmental Psychology, Faculty of Psychology, Universidade de Santiago de Compostela, Santiago de Compostela, A Coruña, Spain

⁴Department of Electronics and Computing, Higher Polytechnic Engineering School, Universidade de Santiago de Compostela, Lugo, Spain

⁵Health Research Institute of Santiago (IDIS), Santiago de Compostela, A Coruña, Spain

⁶Galicía Sur Health Research Institute (IISGS), Vigo, Pontevedra, Spain

*xana.cid.mejias@usc.es



BACKGROUND

Identifying the **neurocognitive profiles** that define the **preclinical** (Subjective Cognitive Decline, SCD) and **prodromal** (Mild Cognitive Impairment, MCI) stages of Alzheimer's disease (AD) is crucial to detect individuals at risk and improve early intervention strategies¹. We developed an optimized model based on neuropsychological measures **to cross-sectionally classify** cognitively unimpaired (CU), SCD, single-domain MCI (sd-MCI), and multi-domain MCI (md-MCI) along the AD continuum, using **machine learning approaches**.

METHODS

We studied 318 participants (138 CU, 73 SCD, 53 sd-MCI, and 54 md-MCI) from the Compostela Aging Study (CompAS²) and used 107 baseline neuropsychological and sociodemographic measures. We compared three single algorithms (Support Vector Machine [SVM], Random Forest [RF], and Logistic Regression [LR]) and a Soft-Voting Ensemble, evaluating performance using ROC curve analysis, and reporting class-wise AUC under leave-group-out cross-validation. Prioritizing early stages, we selected **SVM** based on SCD performance and optimized the feature set to **13 features** using performance-based selection and a collinearity-reduction procedure. The importance of the included features was assessed using permutation importance analysis.

RESULTS

The final SVM model (13 features) showed high discrimination between the diagnostic groups considered, with class-wise **AUC values of 0.82 (CU), 0.83 (SCD), 0.80 (sd-MCI), and 0.86 (md-MCI)** (Figure 1). The greatest performance decreases under permutation were observed for short-delay free recall measure (.190), memory complaints reported by the patient (.067), and age (.058) (Figure 2), highlighting **episodic memory, cognitive complaints, and age** as the most influential variables in differentiating the groups along the AD continuum.

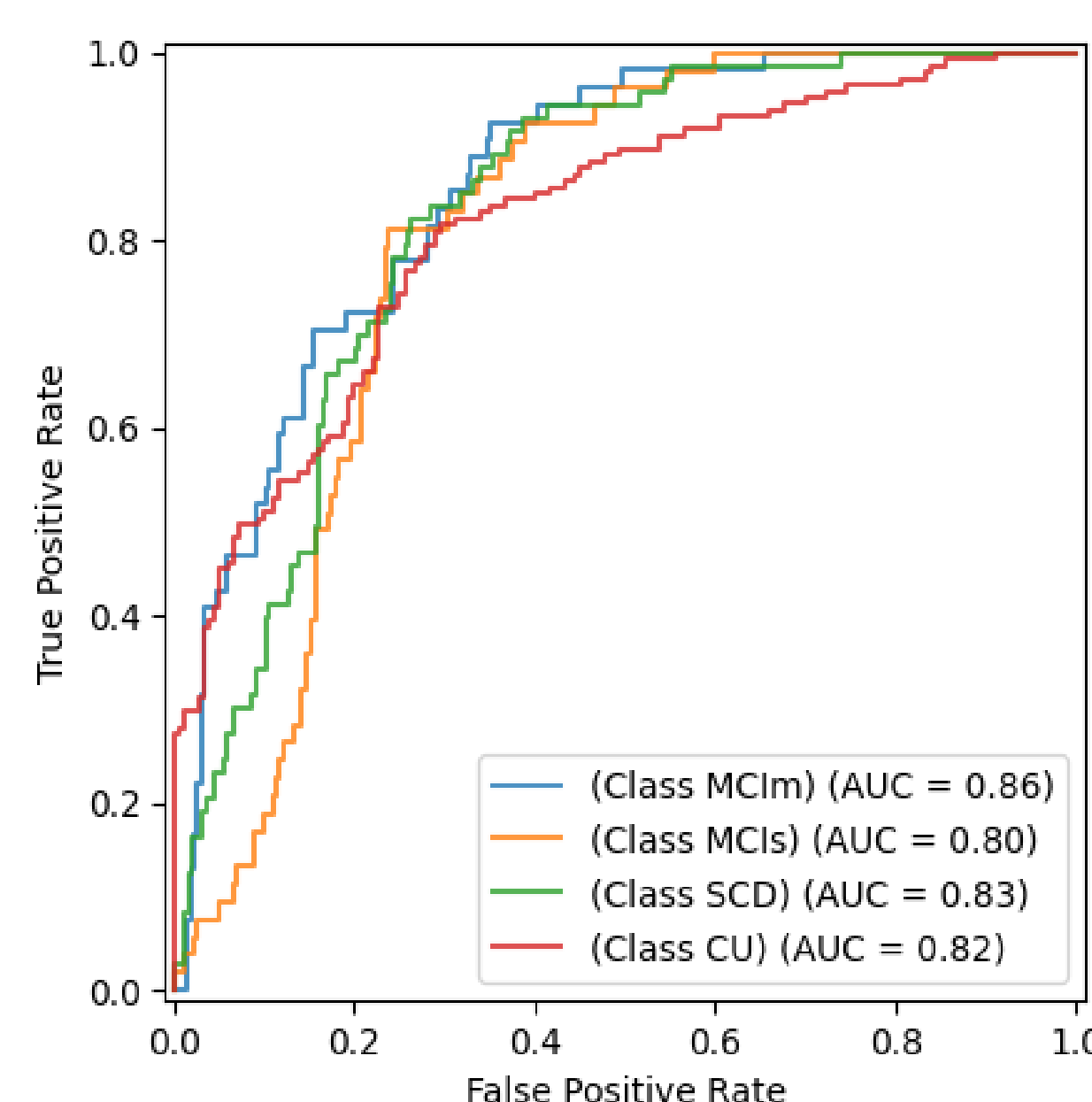


Figure 1. ROC curves for the Neuropsychological SVM Model (13 features), with class-wise AUC values.

Note: AUC= Area Under the Curve; MCI= Multi-domain MCI; MCI= Single-domain MCI; SCD= Subjective Cognitive Decline; CU= Cognitively Unimpaired.

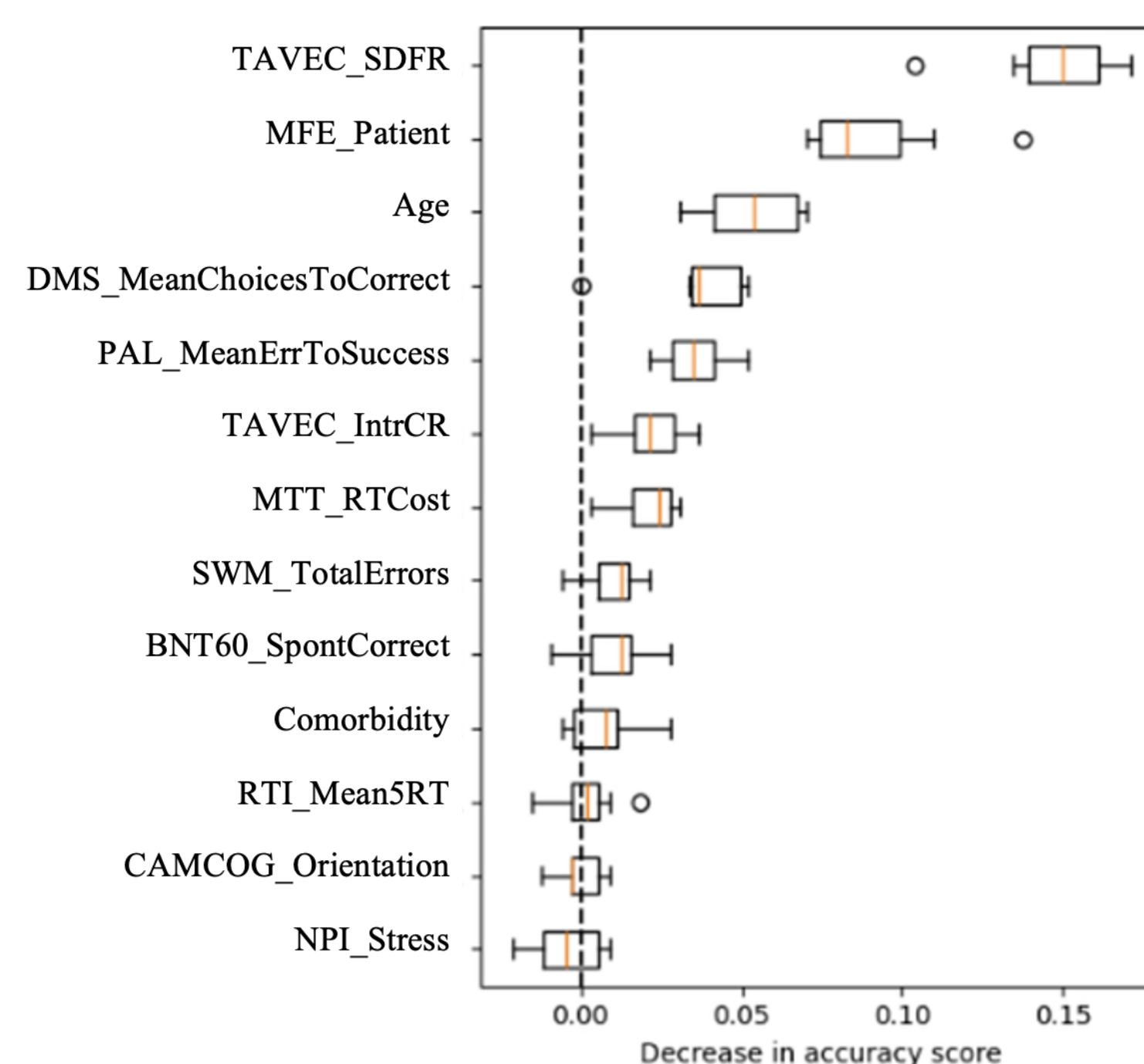


Figure 2. Permutation importance of the 13 features included in the Neuropsychological SVM Model.

Note: TAVEC= Test de Aprendizaje Verbal España-Complutense; MFE_Patient= Memory Failures of Everyday (patient-reported); DMS= Delayed Matching to Sample task from the CANTAB battery; PAL= Paired Associates Learning task from the CANTAB battery; MTT= Multitasking Test task from the CANTAB battery; SWM= Spatial Working Memory task from the CANTAB battery; BNT60= Boston Naming Test, 60-item version; RTI= Reaction Time task from the CANTAB battery; CAMCOG_Orientation= Orientation subscale score from the Cambridge Cognitive Examination-Revised (CAMCOG-R); NPI_Stress= Neuropsychiatric Inventory-Questionnaire (NPI-Q) (caregiver distress total score).

CONCLUSIONS

The neuropsychological SVM model demonstrated good discrimination across the preclinical and prodromal stages of AD, highlighting episodic memory and patient-reported memory complaints as key predictors. These findings suggest that this algorithm **could support clinical practice** by helping to identify individuals at higher risk in the early stages and **guiding future clinical decision-making**.



TAKE-HOME MESSAGE

- ✓ Neurocognitive measures can cross-sectionally discriminate among early stages along the AD continuum.
- ✓ An SVM Model with 13 features showed good discrimination and may guide future clinical decision-making.
- ✓ Episodic memory, patient-reported cognitive complaints, and age emerged as the most relevant predictors.

