Supplementary material

In this supplementary material, we provide a comprehensive overview of the significance of all features (not limited to the five most crucial ones highlighted in the main text) in distinguishing between various cognitive states, utilizing the ANOVA method. We also provide the cumulative AUC-ROC measures of classification in the function of the number of the most significant characteristics. The results are presented in the Figs 1, 2, 3.

As indicated by the presented figures, the feature importance, as measured by the F1 score, decreases exponentially across all cases. Conversely, the cumulative AUC-ROC values fluctuate as more features are included in the classification process. This fluctuation can likely be attributed to overfitting, where an increase in the number of features, while maintaining a constant sample size, causes the model to struggle in identifying the optimal parameters for fitting the high-dimensional data.

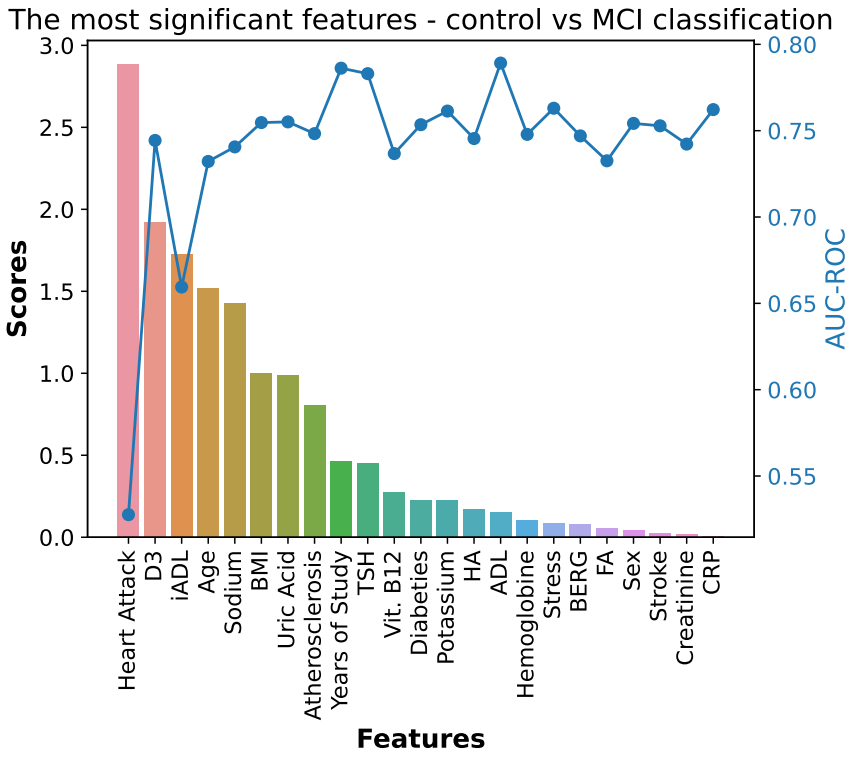

Figure 1. The bar plot displays the most significant features for classifying healthy subjects (control) versus individuals with mild cognitive impairment (MCI), ranked in descending order of importance based on ANOVA test results. "HA" and "FA" denote arterial hypertension and atrial fibrillation, respectively. The blue line illustrates the cumulative AUC-ROC as a function of the number of features included in the classification using a non-linear SVM. For instance, the AUC-ROC at the iADL point on the x-axis represents the value achieved when the top three features—Heart Attack, Vitamin D3, and iADL—are considered in the classification.

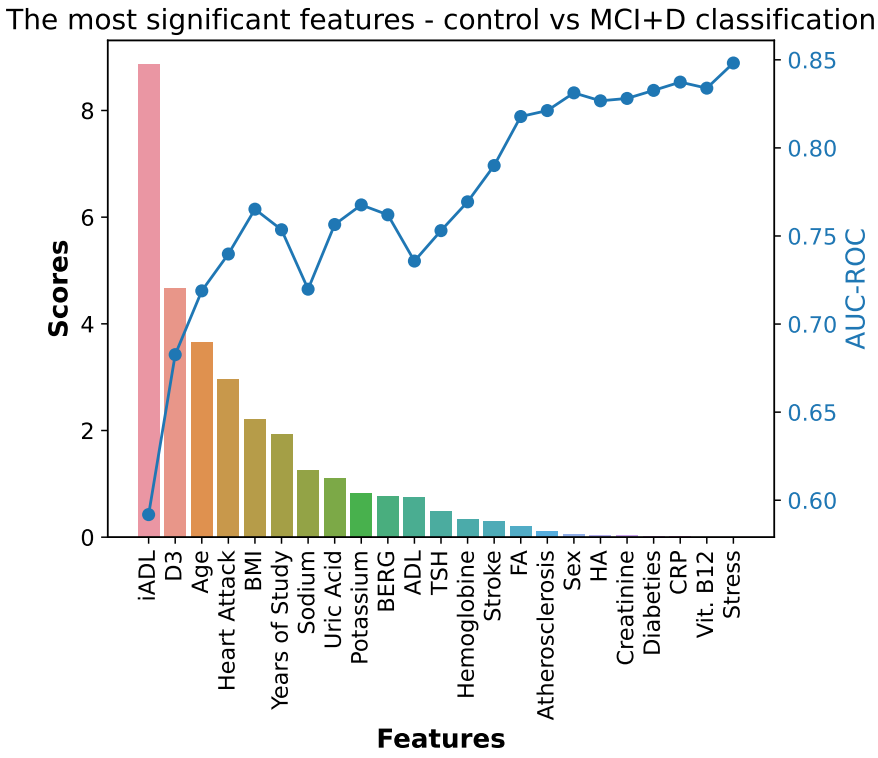

Figure 2. The same as in the caption of the Fig. 1 but for the classification between control and combined group of MCI and dementia subjects with random forest classifier.

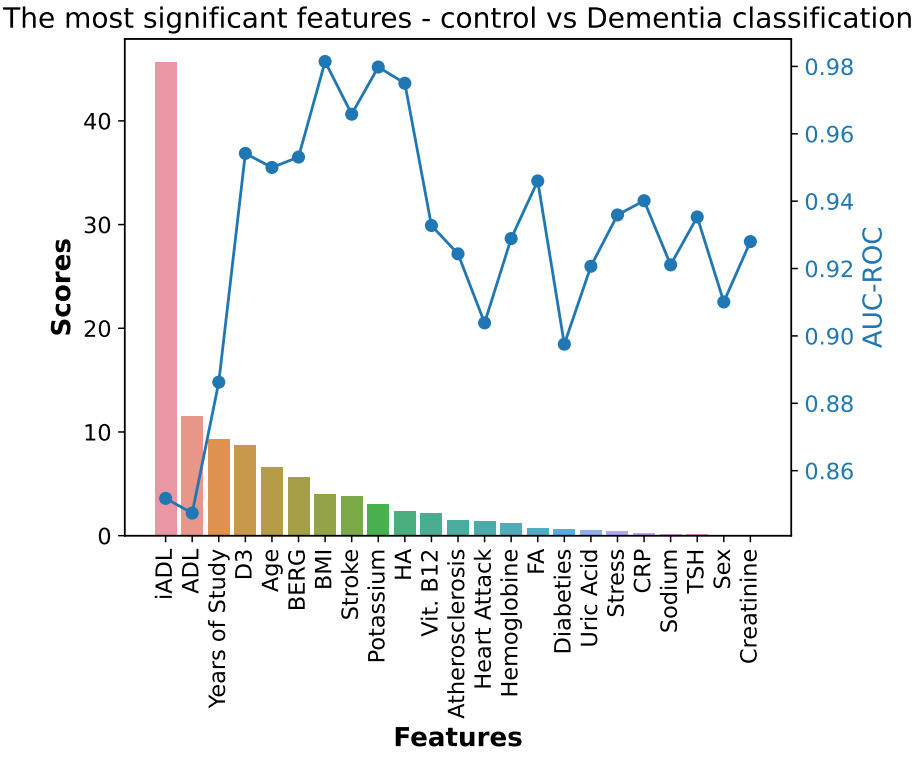

Figure 3. The same as in the caption of the Fig. 1 but for the classification between control and dementia subjects.