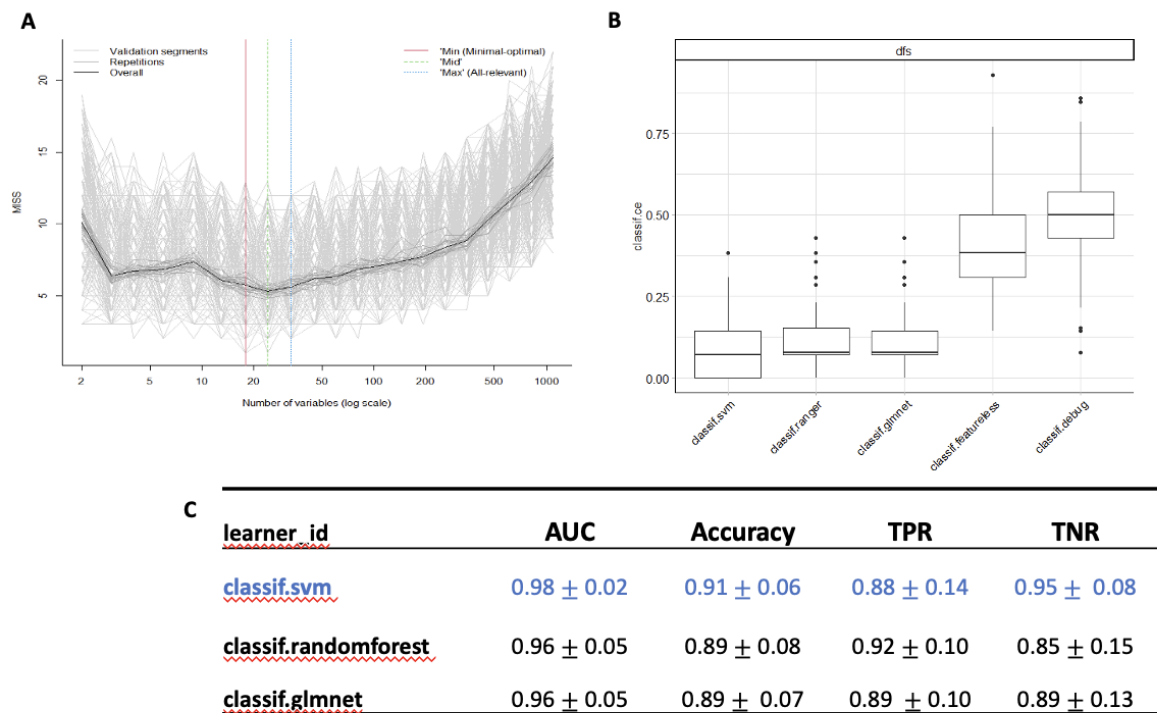
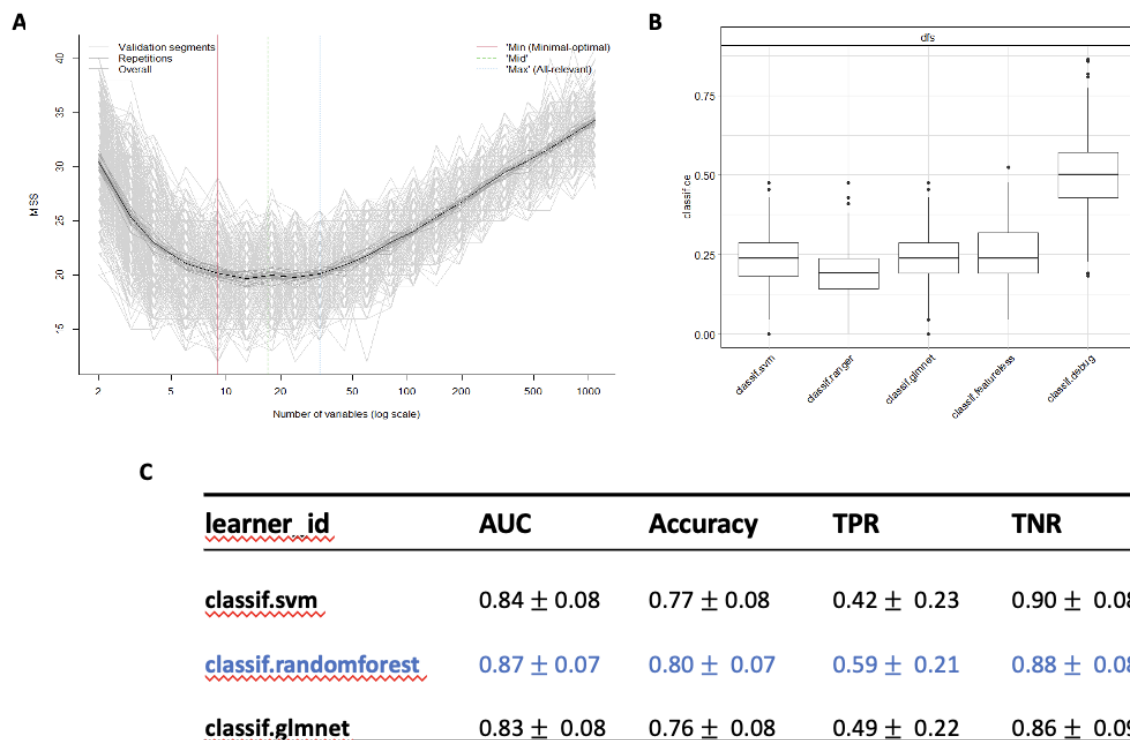


## Supplementary figures:

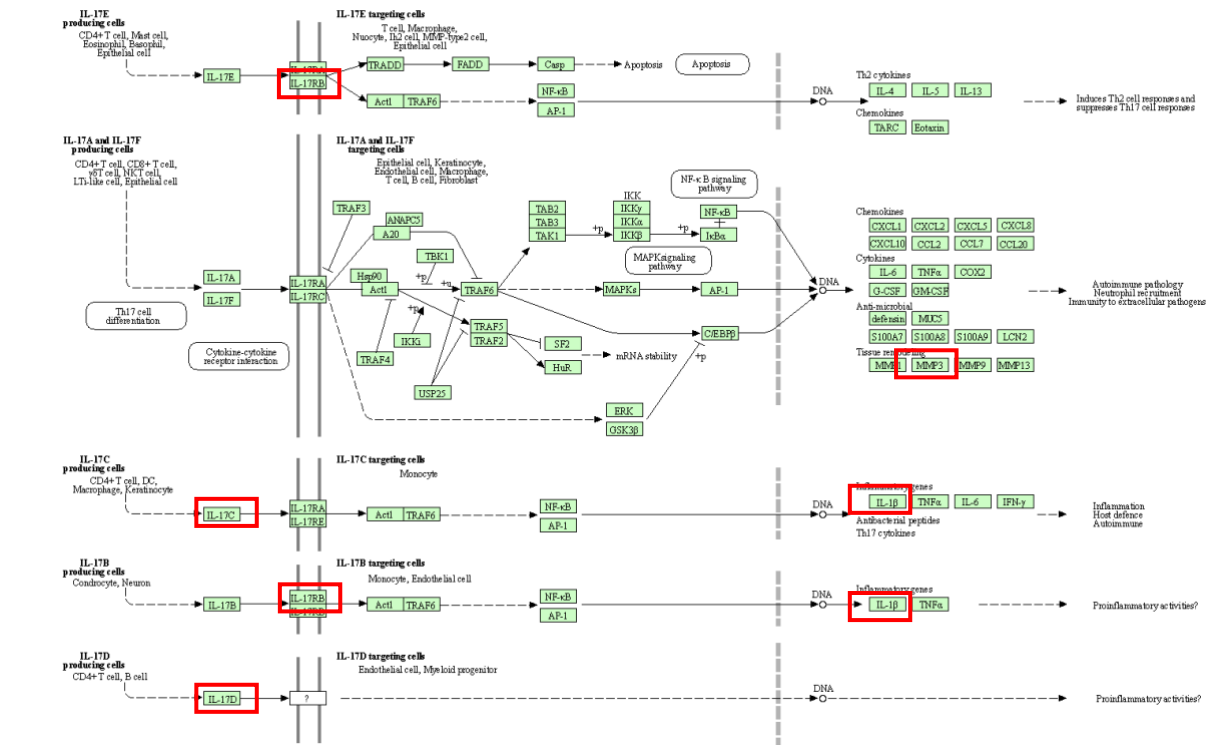


**Figure S1. Machine learning approach for predictors identification and evaluation (dementia vs controls).** **A).** The misclassification rate as a function of the number of included variables. The light grey lines in the graph represent the validation performance of individual inner segments, while the darker grey lines show the average validation performance of all inner segments. The minimal-optimal ("Min") and all-relevant ("Max") models represent the two extremes of variable selection, with validation performance (measured in misclassifications) minimized within 5% of the minimum. The minimal-optimal model represents the minimum number of variables needed for optimal method performance, such as in biomarker discovery, while the all-relevant model includes all variables with relevant signal-to-noise for the research question, including redundant but non-erroneous variables. The "Mid" model represents a compromise between the "Min" and "Max" models, found at the geometric mean, representing the 17 predictors. **B)** A Comparison of classification error rates for five different algorithms: SVM, RF, GLMnet, Featureless, and Random Guessing on the panel of the 17 proteins identified by MUVr and Boruta. The Featureless algorithm assigns all samples to the largest class, while the Random Guessing algorithm randomly assigns samples to classes. These algorithms are included for comparison. The SVM algorithm showed the best performance, with error rates similar to those of the Random Forests and GLMnet algorithms and much better than Featureless and Random Guessing algorithms. **(C)** A comparison table of performance measures of the five ML methods.

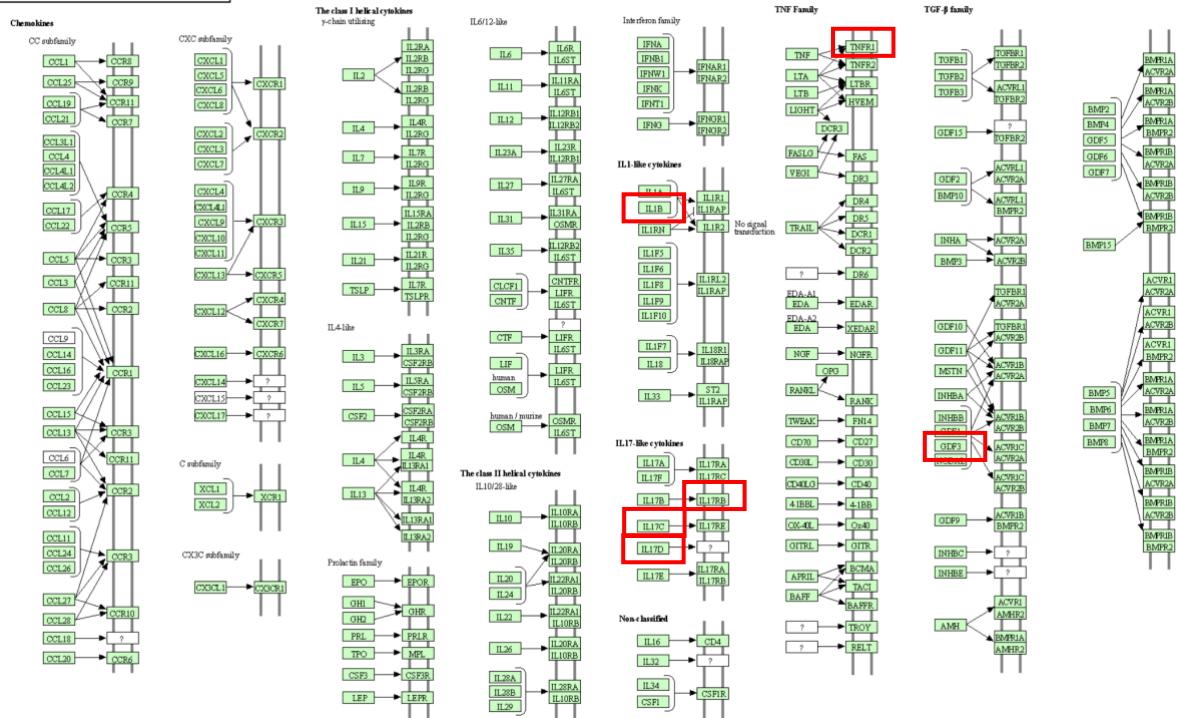


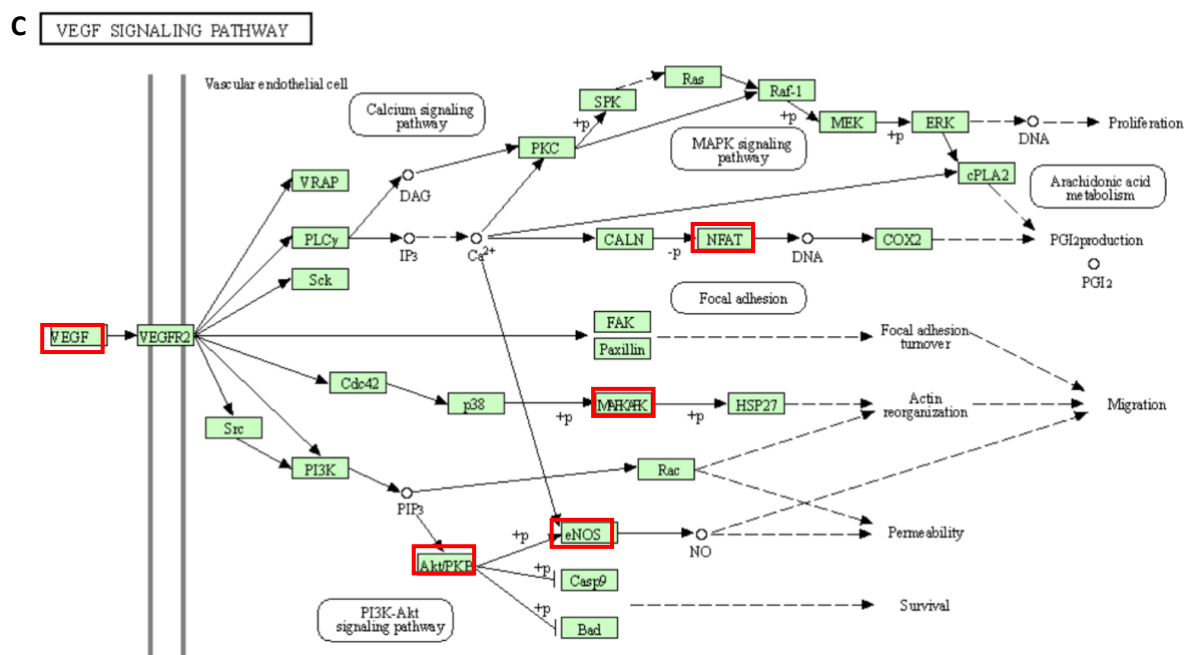
**Figure S2. Machine learning approach for predictors identification and evaluation (dementia vs MCI).** **A)** The MUVr misclassification rate as a function of the number of included variables. The light grey lines in the graph represent the validation performance of individual inner segments, while the darker grey lines show the average validation performance of all inner segments. The minimal-optimal ("Min") and all-relevant ("Max") models represent the two extremes of variable selection, with validation performance (measured in misclassifications) minimized within 5% of the minimum. The minimal-optimal model represents the minimum number of variables needed for optimal method performance, such as in biomarker discovery, while the all-relevant model includes all variables with relevant signal-to-noise for the research question, including redundant but non-erroneous variables. The "Mid" model represents a compromise between the "Min" and "Max" models, found at the geometric mean, representing the 13 predictors. **B)** A Comparison of classification error rates for five different algorithms: SVM, RF, GLMnet, Featureless, and Random Guessing on the panel of the 8 proteins identified by MUVr and Boruta. The Featureless algorithm assigns all samples to the largest class, while the Random Guessing algorithm randomly assigns samples to classes. These algorithms are included for comparison. The random forest algorithm showed the best performance, with error rates similar to those of the SVM and GLMnet algorithms and much better than Featureless and Random Guessing algorithms. **(C)** A comparison table of performance measures of the five ML methods.

**A**

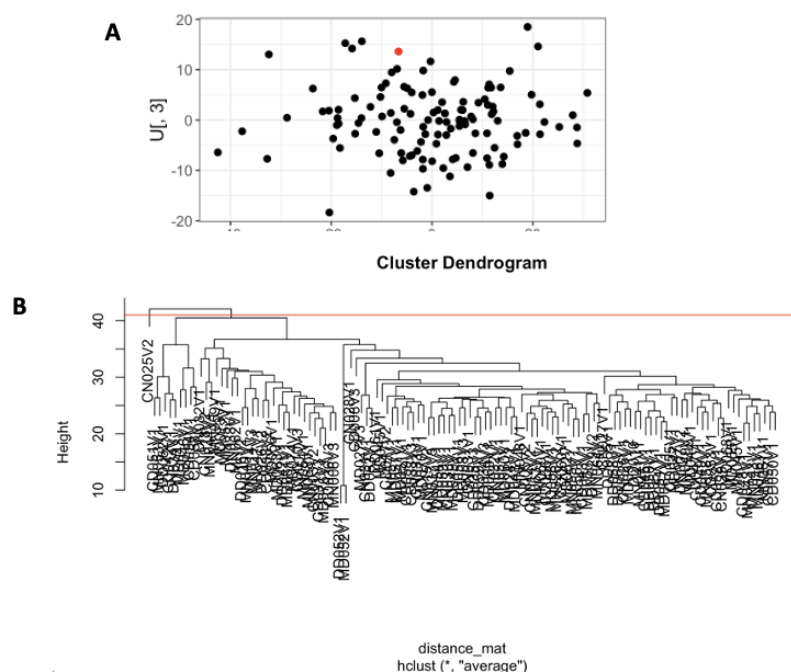


**B**





**Figure S3. Dementia dysregulated pathways. (A)** IL17 signaling pathway. **(B)** Cytokines-cytokines receptors interaction. **(C)** VEGF signaling pathway. In red are the proteins that were differentially expressed in dementia.



**Figure S4. Outliers detection. (A)** One sample was identified as an outlier due to a covariate-adjusted PCA value more than four standard deviations from the centroid. **(B)** The same sample was confirmed by clustering.