

# Incorporating Multi-Omics for Alzheimer’s Disease Predictions

**Author:** James Lee ([chunhonglee@tudelft.nl](mailto:chunhonglee@tudelft.nl))      **Responsible Professor:** Marcel Reinders      **Supervisors:** Timo Verlaan, Roy Largenoije, Gerard Bouland



## Background:

- Alzheimer's Disease (AD) is a complex age-related neurogenerative brain disease, with currently no cure available
- The current leading hypothesis "Amyloid Cascade Hypothesis", states that the increase of the protein Amyloid- $\beta$  activates the tau pathology
- Research has been done on single-omics to better understand this relationship
- Development in multi-omics technologies raises the question, whether using multiple omics, could provide better classifications or gain a deeper insights in the mechanism of the disease

## Research Question:

- Does incorporating different omics, or a combination of multiple omics, improve Alzheimer’s Disease predictions?

## Dataset:

- Religious Orders Study and Memory and Aging Project (ROSMAP)
  - Proteomics (LC-SRM), the set of proteins;
  - Metabolomics (Metabolon HD4), the set of molecules because of metabolism;
  - Epigenetics (ChIP Seq), looks at how cells control gene activity without changing the DNA;
  - Gene Expression (RNA array), measures the levels of mRNA;
- Classify the “Final Consensus Cognitive Diagnosis (cogdx)”
  - The overall cognitive diagnosis, neurologists gave after reviewing the clinical data from a patient, after death

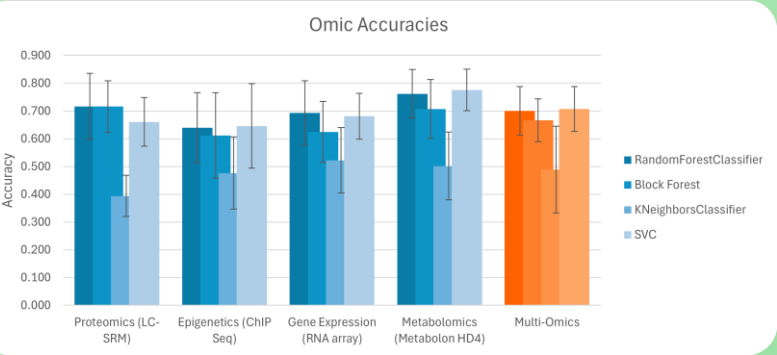
## Methodology:

- Models
  - Random Forest
  - Block Forest
  - $k$ -Nearest Neighbors
  - Support Vector Machines (SVMs)
- Implementation Details
  - 10-Fold Stratified Cross-Validation
  - Mean normalization
  - Oversampling using Synthetic Minority Over-Sampling Technique (SMOTE)
  - Feature selection by calculating the ANOVA F-value

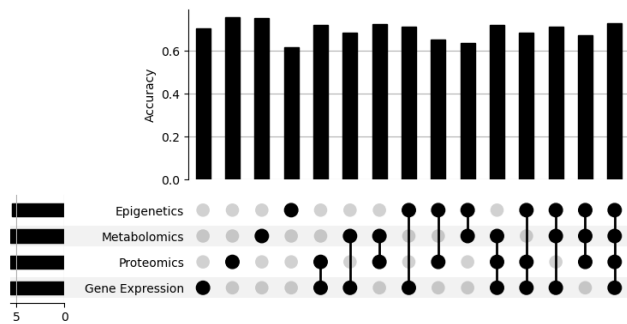
## Conclusion:

- The use of multi-omics did not improve the predictions of Alzheimer’s Disease compared to using single-omics

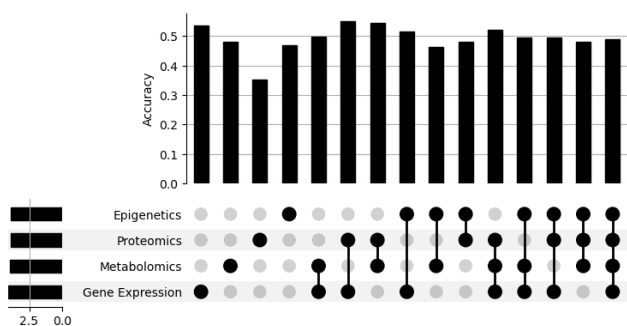
## Results:



### Random Forest



### $k$ -Nearest Neighbors



### Support Vector Machines

