Relating Biological and Clinical Features of Alzheimer‘s Patients With Predictive Clustering Trees

Martin Breskvar[1,2], Bernard Ženko[2], Sašo Džeroski[1,2]

[1]Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
[2]Jožef Stefan Institute, Ljubljana, Slovenia

Human Brain Project

European Commission

Outline

- Motivation and Goals
- Data
- Methodology: Predictive Clustering Trees
- Analysis of Results
- Conclusions

What is Alzheimer‘s Disease?

- Alzheimer‘s disease (AD) is a form of dementia
- A lot of research: Alzheimer‘s Disease Neuroimaging Initiative (ADNI), Human Brain Project, BRAIN Initiative,...
- Some treatment available but AD cannot be stopped; it can only be slowed down to a certain degree

Motivation

- In practice AD diagnosis is mainly based on clinical criteria that can be subjective → poor understanding of the disease
- 47.5 million people affected worldwide

  → expected to triple by 2050
- Total global costs estimated at US$ 604 billion

  → More costly than breast and prostate cancer combined

Identification of connections between biological and clinical variables is a long term goal

- Biomarker identification
- Drug development
- Predicting disease progression
- ...

Goals

- We aim at finding homogeneous subpopulations, where identification of such connections should be easier. We will try to find them by observing the differences in everyday cognition

- According to Farias et. al., everyday cognition shows promise as a tool for measuring general and domain specific everyday functions in the elderly

- We employ a method of Predictive Clustering Trees, which can generate homogeneous subpopulations with respect to clinical and biological variables

Everyday Cognition

- Everyday cognition (ECog) is a questionnaire
- It requires cooperation of patients and their study partners
- It assesses the patient's capability to perform normal, everyday tasks
- Patients and study partners compare the patient's current activity levels with levels from 10 years prior the examination
- The domains of memory, language and executive functioning are assessed
- Answers are evaluated on a 5 point scale:
  (1) no change or performing better,
  (2) occasionally performs worse,
  (3) consistently performs worse,
  (4) performs much worse,
  (5) does not know

Dataset: ADNI

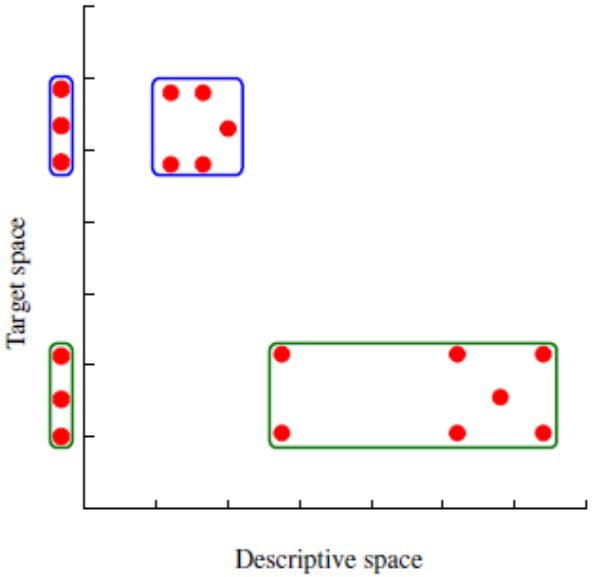
- Alzheimer's Disease Neuroimaging Initiative (ADNI)
- 659 patients: 342 male, 317 female
- 56 biological and 187 clinical attributes at baseline stage
- Biological attributes: ABETA peptides, APOE4 genetic variations, intracerebral volume (ICV), lab measurements (like glucose and protein levels, red and white blood cell counts,…), MRI volumetric data (Ventricles, Hippocampus, WholeBrain, Entorhinal gyrus, Fusiform gyrus, Middle temporal gyrus), TAU and PTAU markers, proteins, and PET imaging results (FDG-PET and AV45).
- Clinical attributes: Alzheimer's Disease Assessment Scale (ADAS13), Mini Mental State Examination (MMSE), Rey Auditory Verbal Learning Test (RAVLT), Functional Assessment Questionnaire (FAQ), Montreal Cognitive Assessment (MOCA), Everyday Cognition, Neuropsychiatric Inventory Examination, Neurological Exam, Modifed Hachinski Ischemia Scale, Geriatric Depression Scale, Baseline symptoms, Clinical Dementia Rating (CDR), Medical History, patient gender and handedness

Machine learning techniques

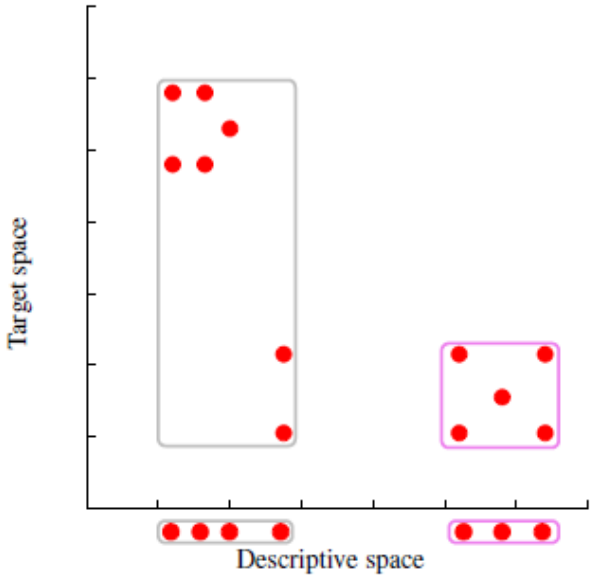
- Predictive modeling: Constructing models that predict a target property of an object from a description of this object
- Clustering: Partitioning objects in clusters of similar objects: High similarity of objects within individual clusters, low similarity between objects from different clusters
- Predictive clustering: combines both; in addition to clusters themselves, we seek for predictive model to be associated with each cluster; predictive model can also predict complex variables

Predictive clustering

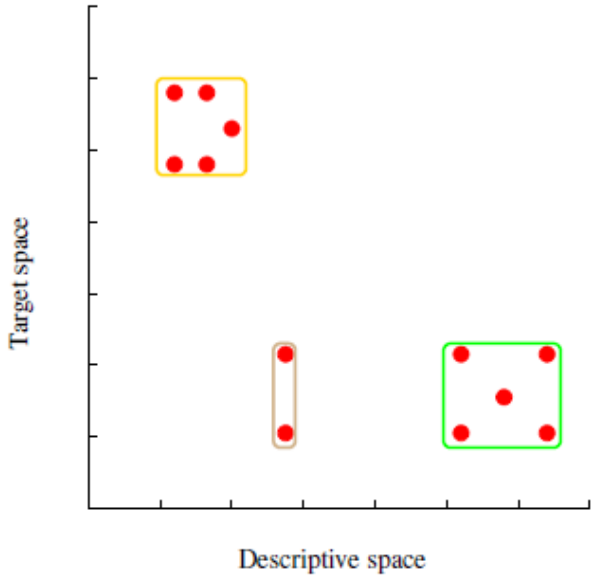
Predictive modeling



Clustering



Predictive Clustering

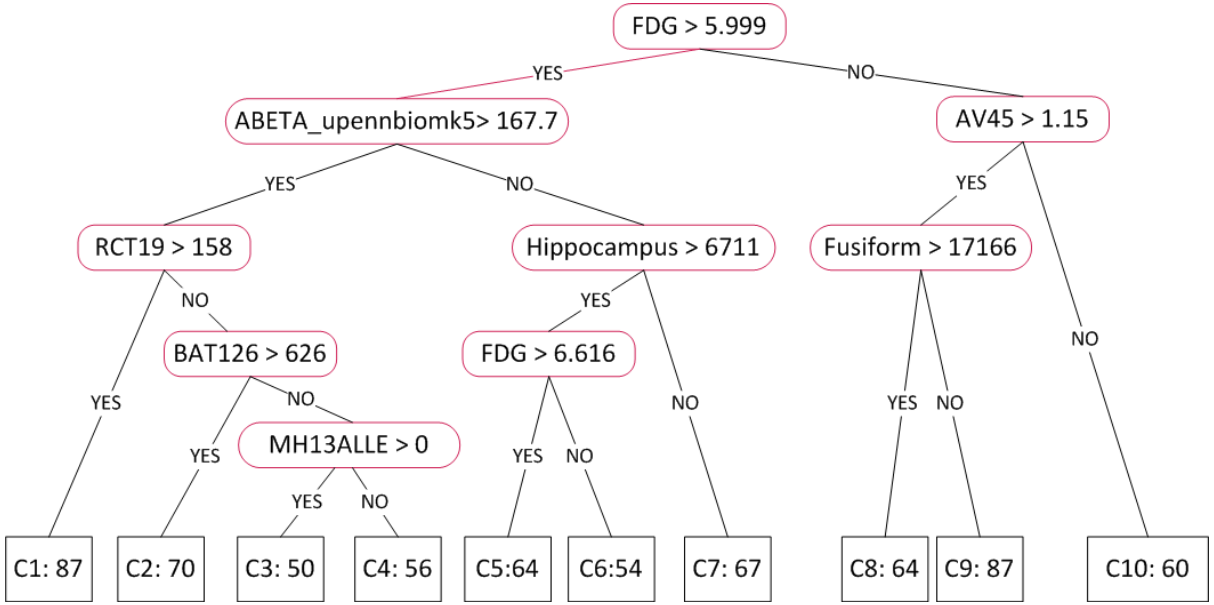


Predictive clustering

- Predictive clustering models are originally implemented as decision trees – Predictive Clustering Trees (PCTs)

- **Result:**
  - Internal structure of the PCT are **interpretable descriptions of** hierarchically ordered **clusters** that contain similar examples (e.g., patients) **in terms of descriptive variables** (e.g., biological variables)
  - that can be used for predicting values of target variables (e.g., everyday cognition)

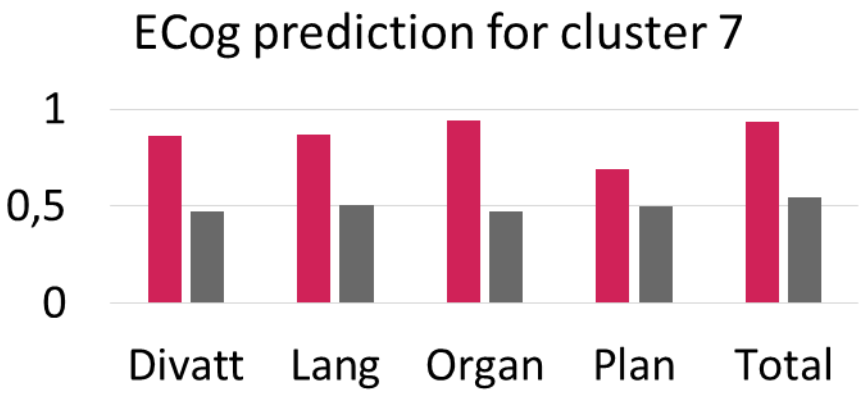
Results

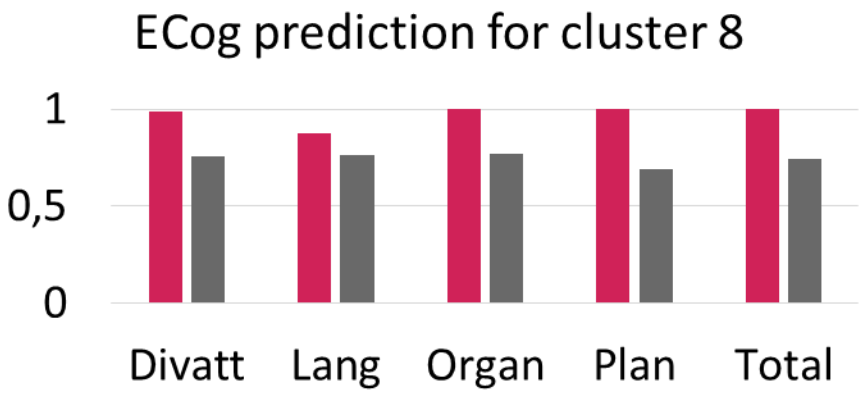


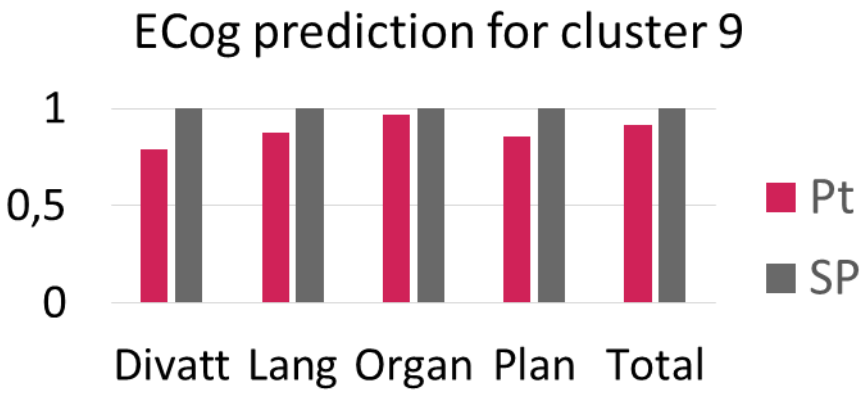
We focusd on clusters where patients are in the late stages of the disease

Cluster assignment	Original diagnosis: CN	SMC	EMCI	LMCI	AD	#ptns
1						87
2						70
3						50
4						56
5						64
6						54
7						67
8						64
9						87
10						60

Results







Results are normalized on the scale from 0 to 1.

Conclusions

- Predictive clustering is a useful tool for detecting clusters of patients that are homogeneous both in respect of their clinical and their biological properties
- The results suggest that in such clusters it is easier to identify relevant relations between clinical and biological properties
- Cluster 9 is characterized by patients with different cognition than clusters 7 and 8, which have similar patient structures.
- Possible explanations:
  - Data collection or processing artefact?
  - Existence of different biological pathways?

Thank you!

Human Brain Project