Hybrid Analysis of Psychiatric Disorders in a Pediatric Dataset

ETH zürich

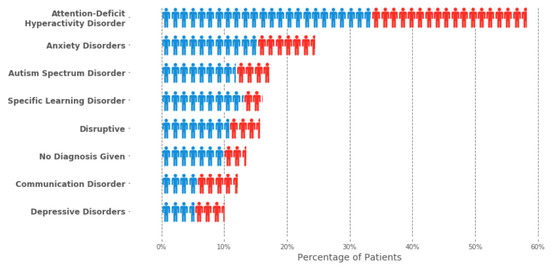
Sotiris Anagnostidis Georgios Vasilakopoulos Adamos Solomou

Department of Computer Science, ETH Zurich

University of Zurich[UZH]

Introduction

- Psychiatric disorders are among the most common and debilitating diseases.
- Diagnosis of psychiatric disorders is highly subjective and involves multiple prolonged interviews conducted by a psychiatrist.
- We develop a data driven framework to aid diagnosis in a reliable and cost-effective manner, overcoming the subjectivity of manual diagnosis.
- **Multimorbidity**: Subjects in our sample are (potentially) diagnosed with multiple disorders.
- Records for each subject include: MRI, EEG data and Behavioural Questionnaires.



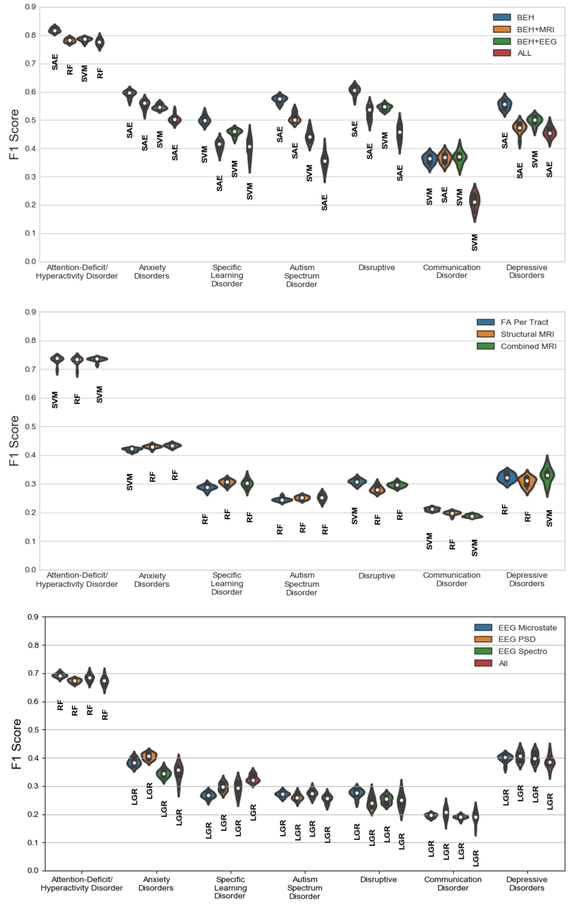
Related Work

- Prior studies focus on differentiating between patients with an isolated psychiatric disorder and "healthy" subjects [1].
- This is in contrast to realistic scenarios, where over 75% of subjects have multiple psychiatric disorders (multimorbidity).

Code & Documentation
sanagno/psychiatry-mri-eeg

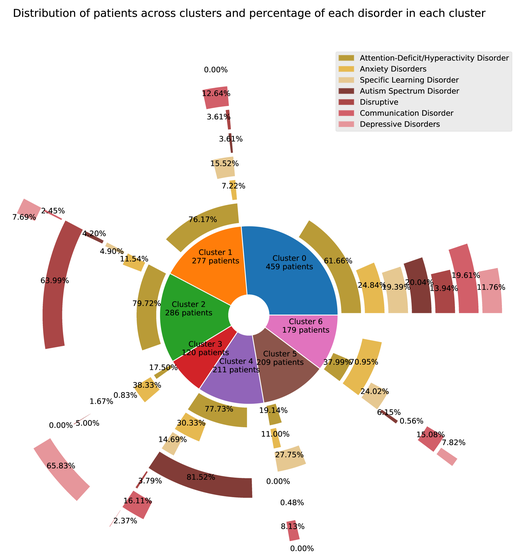
Predicting Psychiatric Multimorbidity

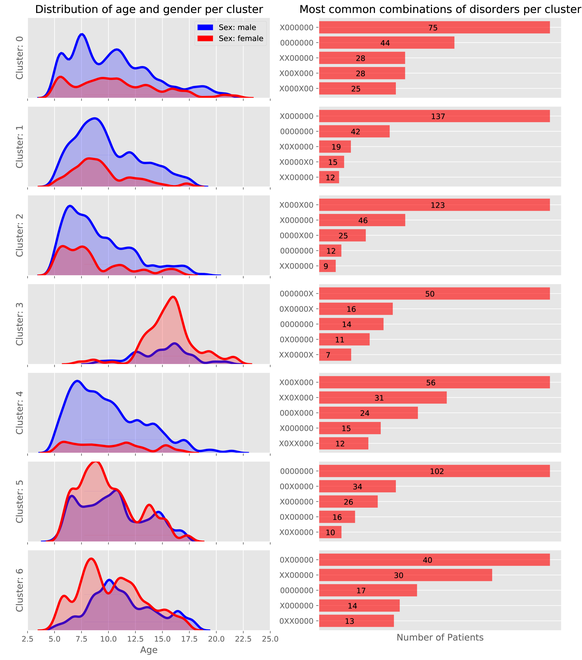
- We decompose the multi-class, multi-label problem into several binary classification tasks.
- Support Vector Machines (**SVM**), Logistic Regression (**LGR**), Random Forest (**RF**).
- Supervised Autoencoder (**SAE**): An autoecoder-based model that jointly learns to predict labels and inputs (reconstruction) [2].
- Performance varies across datasets and models.



Data Driven Clustering

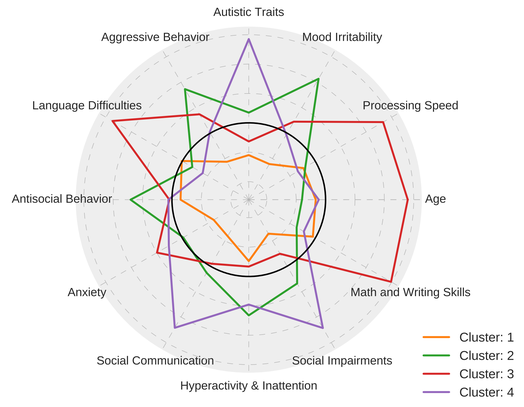
- We employ k-Means clustering to group subjects based on their latent representation (inferred by SAE).





Feature Importance per Cluster

- Identify informative characteristics per cluster.



Predicting Cluster Driven Labels

- We introduce new target categories based on the outcome of our clustering procedure.
- MRI & EEG data are used to predict the devised labels.
- To allow comparison, we employ a hierarchical clustering process to group subjects based on their label annotation.

Target	Balanced Accuracy
Data Driven Clusters	0.2653 ± 0.0529
Label Driven Clusters	0.2054 ± 0.0424

[1] Choong-Wan Woo, Luke Chang, Martin Lindquist, and Tor Wager. Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, 20:365–377, 02 2017.

[2] Lei Le, Andrew Patterson, and Martha White. Supervised autoencoders: Improving generalization performance with unsupervised regularizers. In *Advances in Neural Information Processing Systems 31*. 2018.