Learning the progression and clinical subtypes of Alzheimer's disease from longitudinal clinical data

Vipul Satone[1], Rachneet Kaur[1], Faraz Faghri[1,2], Mike A Nalls[2], Andrew B Singleton[2], Roy H Campbell[1]

[1] University of Illinois at Urbana-Champaign, USA
[2] National Institutes of Health, USA

Goal

Identify Alzheimer's disease (AD) progression space and subtypes through time variant clinical data.

Method

- Performed data preprocessing, cleaning, imputation, normalization, and vectorization
- Used Non-negative matrix factorization (NMF) to construct progression space
- Used Gaussian mixture model (GMM) to identify different clusters in the projected space
- Developed predictive model using Random forest for progression stage after 24 and 48 months from baseline

Results

- Developed "Alzheimer's disease progression space"
- Interpretable progression space dimensions: memory decline and cognitive decline.
- Identified three distinct progression clusters corresponding to low moderate and high progression rate of AD
- Predicting AD progression rate after 24 and 48 months from baseline with an average AUC of 0.93

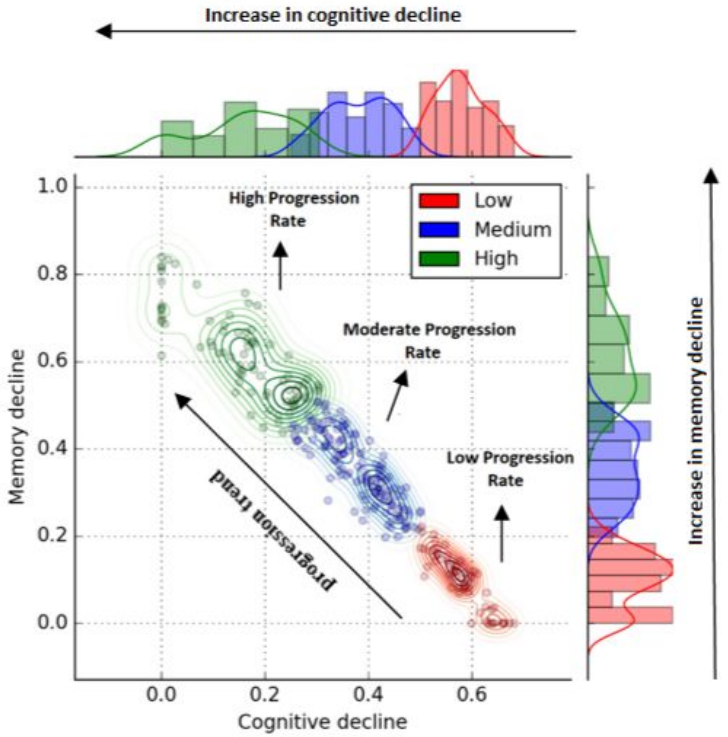

Figure: Progression rates classified by Gaussian mixture model. Plots on x-axis and y-axis shows density variation within each cluster