Bern Data Science Day 2021 - April 23

- **Title:**
  Sequential neutral-zone classification for diagnosis of dementia from longitudinal measures with mixed-effects models.
- **Topic (from Webpage):**
  Data Science applications and challenges in Medicine, Natural Sciences, and Engineering
- **Name and Email:**
  Patric Wyss (PhD student in Biomedical Sciences at GCB Bern)
  patric.wyss@upd.unibe.ch
- **Affiliation**:
  University Hospital of Old Age Psychiatry and Psychotherapy, University of Bern

Abstract (423 words)

Repeated measurements of biological or cognitive markers of neurodegeneration over time are predictive for the diagnosis of Alzheimer's disease. Clinical applicability of computer assisted prognostic tools require handling of irregular longitudinal data with varying number of observations. Mixed-effects model based generative classification algorithms are suitable to implement these aspects. In this study we (1) introduce performance measures in the context of Alzheimer's disease classification to identify suitable mixed-effects model structures and (2) implement an automatic, sequential algorithm with the flexibility adding measurements until enough evidence for a final decision is incorporated.

In this study we aimed at separating patients with mild cognitive impairments that do (MCI-converter) or do not (MCI-non-converter) convert to the stadium of Alzheimer's disease based on longitudinal measurements of hippocampal volume or cognitive scores. We derived neutral zone classifiers which add a no-decision option as a classification outcome. Neutral zone classification was either defined by the allowed false positive and negative rates (using p-values of classification) or decision costs. We quantitively compared different mixed-effects models to infer about underlying structures (in terms of accuracy, AUC, or measures from neutral zone classification). In addition, we implemented sequential decision strategies that either choose one of the classes or collect new information until some criterion is fulfilled or there is no observation left for a subject. We compared our adaptive strategy that considers expected future costs to an already existing sequential approach where recommendations are made solely based on costs.

Our results show that a decomposition of longitudinal and cross-sectional effects (fixed effects for baseline age and time since baseline) as well as the inclusion of subject-specific deviations from the longitudinal trends of the population the subject belongs to (random effects for time per subject) leads to better classification performances. This was true for both biological and cognitive markers. Moreover, in contrast to the descriptive sequential approach that does not consider the expected cost reduction when assigning the label neutral, the adaptive approach using the generative information in the models have low costs by deciding earlier for one of the classes.

In this study we found new evidence about the underlying structure of longitudinal assessed markers of neurodegeneration and present a paradigm that allows to fine tune mixed-effects models structure based on clinically relevant criteria in dementia diagnosis. Moreover, we showed how mixed-effects model based generative classification frameworks can be used to build sequential decision strategies allowing to optimally balance costs and performance in the diagnosis of dementia. This work lays the foundation for multivariate adaptive and multiple-stage sequential decisions with mixed-effects models.