Learning transition times in event sequences: the Event-Based Hidden Markov Model

Connecting ideas from event-based and hidden Markov modelling to derive a new interpretable model of disease progression

Progressive diseases such as Alzheimer’s disease (AD) are characterised by monotonic deterioration in functional, cognitive and physical abilities over a period of years to decades. Data-driven models of disease progression can be used to learn hidden information, such as individual-level stage, from observed data. Here we address the problem of how to learn transition times in event sequences of disease progression, by introducing a new generative Event-Based Hidden Markov Model (EB-HMM) of disease progression. The main novelties of our work are as follows.

• We generalise a formerly cross-sectional model (the EBM: event-based model Fonteijn et al. 2012), allowing it to account for longitudinal data.
• We define a Bayesian ‘event-based’ framework to inject prior information into structured inference from longitudinal data.
• We use our model to learn a new clinically interpretable sequence and timing of events in AD and to predict individual-level trajectories.

Results

Inferring the timeline of feature changes in Alzheimer’s disease

We use data from the ADNI study, a longitudinal multi-centre observational study of AD. We select 468 participants (119 CN: cognitively normal; 297 MCI: mild cognitive impairment; 29 AD: manifest AD; 23 NA: not available), and three time-points per participant (baseline and follow-ups at 12 and 24 months). Individuals were allowed to have missing data at any time-point.

Methods

The Event-Based Hidden Markov Model

To formulate EB-HMM, we make three assumptions, namely i) monotonic feature changes; ii) a consistent event sequence, S, across the whole sample; and iii) Markov (memoryless) stage transitions.

Assuming independence between observed features i=1,...,I, if a patient j=1,...,J is at latent state k=0,...,N at time t=1,...,T in the progression model, the likelihood of their data Yj is given by:

Here θp = [μp,σp,wp] and θc = [μc,σc,wc] are the mean, μ, standard deviation, σ, and mixture weights, w, for the patient and control mixture model distributions, which define the event-based model. For a full derivation see Wijeratne & Alexander, 2020.

Results

Inferring the timeline of feature changes in Alzheimer’s disease

We use data from the ADNI study, a longitudinal multi-centre observational study of AD. We select 468 participants (119 CN: cognitively normal; 297 MCI: mild cognitive impairment; 29 AD: manifest AD; 23 NA: not available), and three time-points per participant (baseline and follow-ups at 12 and 24 months). Individuals were allowed to have missing data at any time-point.

Results

Inferring the timeline of feature changes in Alzheimer’s disease

We use data from the ADNI study, a longitudinal multi-centre observational study of AD. We select 468 participants (119 CN: cognitively normal; 297 MCI: mild cognitive impairment; 29 AD: manifest AD; 23 NA: not available), and three time-points per participant (baseline and follow-ups at 12 and 24 months). Individuals were allowed to have missing data at any time-point.

Results

Inferring the timeline of feature changes in Alzheimer’s disease

We use data from the ADNI study, a longitudinal multi-centre observational study of AD. We select 468 participants (119 CN: cognitively normal; 297 MCI: mild cognitive impairment; 29 AD: manifest AD; 23 NA: not available), and three time-points per participant (baseline and follow-ups at 12 and 24 months). Individuals were allowed to have missing data at any time-point.

Discussion

A new interpretable model of disease progression

A key corollary benefit of EB-HMM’s formulation is that it can infer probabilistic estimates of group- and individual-level progression from datasets with missing data, both in terms of observed features and time-points. This gives EB-HMM high utility in clinical applications where resources are scarce and it is too expensive to observe a patient multiple times, making EB-HMM an ideal tool for advancing on the objective of accessible healthcare. Future work with EB-HMM will be focused on relaxing its assumptions, in particular allowing for non-monotonic trajectories and multiple event sequences (subtypes).

References
