Contribution ID: **115** Type: **not specified**

Semi-Supervised PARAFAC2 Decomposition for Computational Phenotyping Using Electronic Health Records

Computational phenotyping uses data mining methods to extract clusters of clinical descriptors, known as phenotypes, from electronic health records (EHR). Tensor factorization methods are very effective in extracting meaningful patterns and have become popular in computational phenotyping. Nevertheless, these techniques mainly focus on regular tensors and are used in a fully unsupervised manner. EHR data is often represented by irregular tensors (due to varying hospital visits) and contains some label information that is useful in extracting meaningful phenotypes. While techniques to decompose irregular tensors in an unsupervised manner have been developed, methods to integrate label information in such decomposition are lacking. In this work, we propose a semi-supervised PARAFAC2 decomposition model to extract meaningful patterns from irregular EHR data by incorporating label information available for a subset of instances. Experiments on a synthetic data set and a case study based on MIMIC-IV data show the superiority of our approach over the benchmarks. This novel computational phenotyping method can potentially facilitate medical decision-making in many healthcare applications.

Special/ Invited session

Classification

Both methodology and application

Keywords

Computational phenotyping; Tensor Decomposition

Primary author: Dr KONYAR, Elif (Georgia Tech)

Co-author: REISI GAHROOEI, Mostafa (University of Florida)

Presenter: REISI GAHROOEI, Mostafa (University of Florida)

Track Classification: AI: Machine Learning and Predictive Analytics