Computational Phenotyping from EHR Data for Predictive Analytics

Dr William CHEUNG
Head and Associate Professor of Computer Science, Hong Kong Baptist University, HKSAR

Computational phenotyping refers to the process of extracting and relevant phenotype from heterogeneous data sources for predict clinically important phenotypes before they are exhibited. In this talk, I will present our recent research work on developing machine learning techniques for computational phenotyping based on electronic health records and medical ontologies. The methods that we developed include: (1) hidden interaction tensor factorization (HITF) which can extract more accurate phenotypes by learning the diagnosis-medication correspondence and the underlying phenotypes simultaneously; (2) Collective Non-negative Tensor Factorization (CNTF) which puts the temporal dimension into consideration with recurrent neural network based regularization to extract phenotypes and at the same time dynamic patient representations to represent their evolving health conditions; (3) Medical Concepts Embedding with Multiple Ontological Representations (MMORE) which is a deep learning model which integrates EHR data and medical ontologies for predictive phenotyping via representation learning and an attention mechanism. For performance evaluation, we applied the methods we developed to the publicly available MIMIC III datasets and show that they outperform a number of state-of-the-art methods for various prediction tasks in healthcare.