

DEPARTMENT OF COMPUTER SCIENCE

MPhil Degree Oral Presentation

MPhil Candidate:	Ms. Lihong SONG
Date	9 August 2019 (Friday)
Time:	2:30 pm - 4:30 pm (35 mins presentation and 15 mins Q & A)
Venue:	FSC703, Fong Shu Chuen Library, HSH Campus

“Medical Concept Embedding with Ontological Representations”

Abstract

Learning representations of medical concepts from the Electronic Health Records (EHRs) has been shown effective for predictive analytics in healthcare. The learned representations are expected to preserve the semantic meanings of different medical concepts, which can be treated as features and thus benefit a variety of applications. Medical ontologies have also been explored to integrate with the EHR data to further enhance the accuracy of various prediction tasks in healthcare. Most of the existing works assume that medical concepts under the same ontological category should share similar representations, which however does not always hold. In particular, the categorizations in the categorical medical ontologies were established with various factors being considered. Medical concepts even under the same ontological category may not follow similar occurrence patterns in the EHR data, leading to contradicting objectives for the representation learning. In addition, these existing works merely utilize the categorical ontologies. Actually, it has been noticed that ontologies containing multiple types of relations are also available. However, rare studies make use of the diverse types of medical ontologies.

In this thesis research, we propose three novel representation learning models for integrating the EHR data and medical ontologies for predictive analytics. To improve the interpretability and alleviate the conflicting objective issue between the EHR data and medical ontologies, we propose techniques to learn medical concepts embeddings with multiple ontological representations. To reduce the reliance on labeled data, we treat the co-occurrence statistics of clinical events as additional training signals, which help us learn good representations even with few labeled data. To leverage the various domain knowledge, we also consider multiple medical ontologies (CCS, ATC and SNOMED-CT) and propose corresponding attention mechanisms so as to take the best advantage of the medical ontologies with better interpretability. Our proposed models can achieve the final medical concept representations which align better with the EHR data. We conduct extensive experiments, and our empirical results prove the effectiveness of the proposed methods.

***** ALL INTERESTED ARE WELCOME *****