Colloquium

Extracting medically interpretable concepts from complex health data

Joyce Ho
University of Texas at Austin

Abstract: Electronic health records (EHRs) are an increasingly important source of patient information. Efficient analysis of EHRs can help address many healthcare problems by improving clinical decisions, facilitating knowledge discoveries, and enabling the development of cost-effective treatment and management programs. However, EHRs pose many formidable challenges for traditional analytics. The data are collected across diverse populations, consist of heterogeneous and noisy information, and have varying time resolutions. Moreover, healthcare professionals are unaccustomed to interpreting high-dimensional EHRs; they prefer concise medical concepts. Thus, a major question is how to transform EHR into meaningful concepts with modest levels of expert guidance.

In this talk, I will discuss two approaches to extract concise, meaningful concepts from certain types of health datasets. First, I will describe a dynamic time series model that tracks a patient's cardiac arrest risk based on physiological measurements (i.e., heart rate, blood pressure, etc.) in an intensive care unit. Our algorithm is inspired by financial econometric and yields interpretability and predictability of a cardiac arrest event. Next, I will present a sparse, nonnegative tensor factorization model to obtain multiple medical concepts with minimal human supervision. Tensor factorization utilizes information in the multiway structure to derive concise latent factors even with limited observations. We applied tensor analysis to real EHRs from the Geisinger Health System to automatically identify relevant medical concepts. Both our models are powerful and data-driven approaches to extract medically interpretable concepts from complex health data.

Tuesday, March 3, 2015, 4:00 pm
Mathematics and Science Center: W303

Mathematics and Computer Science
Emory University