University of Pittsburgh | School of Computing and Information

Proposal Defense

Doctor of Philosophy in Intelligent Systems

"**Exploring Data and Clinical Knowledge Driven Computer-Aided Sequential Diagnosis System in Emergency Department**" by **Diyang Xue**

Date: April 26, 2021
Time: 10:00AM – 12:00PM
Place: https://pitt.co1.qualtrics.com/jfe/form/SV_9FSE4NGZL3GdTcG

Committee:

- Dr. Daqing He, Professor, Department of informatics and networked systems, School of Computing and Information
- Dr. Gregory F. Cooper, Professor, Department of Biomedical Informatics, School of Medicine
- Dr. Michael M. Wagner, Professor, Department of Biomedical Informatics, School of Medicine
- Dr. Adam N. Frisch, Assistant Professor, Department of Emergency Medicine, School of Medicine

Abstract:
Medical diagnosis is the process of determining the nature of a disease and distinguishing it from other similar diseases. A diagnosis error happens when a diagnosis is missed, inappropriately delayed, or is wrong. Diagnosis error accounts for the most severe patient harm, the largest fraction of claims, and highest total penalty payouts. One way to reduce diagnostic error is to use a computer-aided diagnostic (CAD) system to augment doctors' diagnostic abilities.
CAD systems can be roughly divided into two types: knowledge-driven systems and data-driven systems. Knowledge-driven systems strongly depend on pre-defined clinical knowledge. The usage of these systems may be restricted by the system's knowledge base; they can only handle diseases and symptoms in the defined knowledge base. Extending and updating these systems is often labor extensive. Data-driven systems learn diagnosis rules from data directly, they can extend to new diseases very easily. Data-driven systems have three main drawbacks: black-box problems, not clinically meaningful, and not clinically actionable.
Few studies have been conducted on how to combine medical knowledge and machine learning algorithms to make a machine learning model clinically meaningful and actionable. We propose to (1) develop a framework that can integrate pre-defined medical knowledge with disease patterns in electronic medical records for disease differential diagnosis; (2) develop an algorithm that generates medical decision trees that recommend diagnosing actions by considering clinical workflow, diagnosis accuracy, and misdiagnosis costs. We plan to evaluate the performance of the developed system in real-world heart disease diagnosis tasks and compare its performance with other multi-label decision trees. We will also exam the effects of medical knowledge and electronic medical records on diagnosis accuracy, misdiagnosis costs, clinical meaningfulness, and actionability.