Announces the Ph.D. Dissertation Defense of

Shuwen Wang

for the degree of Doctor of Philosophy (Ph.D.)

“Tackling Bias, Privacy, and Scarcity in Health Data Analytics”

October 16, 2023, Time 2:30 PM – 4:00 PM
Building, Room EE 405
777 Glades Road
Boca Raton, FL

DEPARTMENT:
Department of Electrical Engineering and Computer Science

ADVISOR:
Xingquan Zhu, Ph.D.

PH.D. SUPERVISORY COMMITTEE:
Xingquan Zhu, Ph.D., Chair
Ankur Agarwal, Ph.D.
Raquel Assis, Ph.D.
Oge Marques, Ph.D.
Xiangnan Zhong, Ph.D.

ABSTRACT OF DISSERTATION
Tackling Bias, Privacy, and Scarcity in Health Data Analytics

Health data analytics has emerged as a critical domain with immense potential to revolutionize healthcare delivery, disease management, and medical research. However, it is confronted by formidable challenges, including sample bias, data privacy concerns, and the cost and scarcity of labeled data. These challenges collectively impede the development of accurate and robust machine learning models for various healthcare applications, from disease diagnosis to treatment recommendations. Sample bias and specificity refer to the inherent challenges in working with health datasets that may not be representative of the broader population or may exhibit disparities in their distributions. These biases can significantly impact the generalizability and effectiveness of machine learning models in healthcare, potentially leading to suboptimal outcomes for certain patient groups. Data privacy and locality are paramount concerns in the era of digital health records and wearable devices. The need to protect sensitive patient information while still extracting valuable insights from these data sources poses a delicate balancing act. Moreover, the geographic and jurisdictional differences in data regulations further complicate the use of health data in a global context. Label cost and scarcity pertain to the often labor-intensive and expensive process of obtaining ground-truth labels for supervised learning tasks in healthcare. The limited availability of labeled data can hinder the development and deployment of machine learning models, particularly in specialized medical domains. This dissertation mainly focuses on health data analytics and explores approaches to tackle the above challenges. More specifically, the following three problems will be studied from different perspectives: (1) Sample bias and specificity in health data. (2) Data privacy and locality in health data. (3) Label cost and scarcity in health data.

BIOGRAPHICAL SKETCH
Born in China
B.S., China Jiliang University, Hangzhou, Zhejiang Province, China, 2019
Ph.D., Florida Atlantic University, Boca Raton, Florida, 2023

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION
Time in Preparation: 2020 - 2023
Published Papers:


Paper under review: