COLLEGE OF ENGINEERING AND COMPUTER SCIENCE
FLORIDA ATLANTIC UNIVERSITY

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

Qualifying Examination Passed: Spring 2021

Published Papers:

1. Shuwen Wang, Magdalyn E. Elkin, and Xingquan Zhu. "Imbalanced learning for hospital readmission prediction using national readmission database." In 2020 IEEE International Conference on Knowledge Graph (ICKG), pp. 116-122. IEEE, 2020.

2. Shuwen Wang, and Xingquan Zhu. "Predictive modeling of hospital readmission: challenges and solutions." IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 5 (2021): 2975-2995.

3. Shuwen Wang, and Xingquan Zhu. "Nationwide hospital admission data statistics and disease-specific 30-day readmission prediction." Health Information Science and Systems, vol. 10, no. 1 (2022): 25.

4. Shuwen Wang, Xingquan Zhu, Weiping Ding, and Amir Alipour Yengejeh. "Cyberbullying and cyberviolence detection: A triangular user-activity-content view." IEEE/CAA Journal of Automatica Sinica, vol. 9, no. 8 (2022): 1384-1405.

5. Shuwen Wang, and Xingquan Zhu. "FedDNA: Federated learning using dynamic node alignment." PLOS ONE, 18, no. 7 (2023): e0288157.

6. Man Wu, Shuwen Wang, Shirui Pan, Andrew C. Terentis, John Strasswimmer, and Xingquan Zhu. "Deep learning data augmentation for Raman spectroscopy cancer tissue classification." Scientific Reports, 11, no. 1 (2021): 23842.

7. Divya Gangwani, Qianxin Liang, Shuwen Wang, and Xingquan Zhu. "An empirical study of deep learning frameworks for melanoma cancer detection using transfer learning and data augmentation." In 2021 IEEE International Conference on Big Knowledge (ICBK), pp. 38-45. IEEE, 2021.

Paper under review:

1. Shuwen Wang, and Xingquan Zhu. "LG-FAL: Locality-customized GSA Federated Active Learning." Under Review.