The College of Graduate Studies and the College of Information Technology

Cordially Invite You to a

PhD Dissertation Defense

Entitled

A NOVEL MULTI-MODEL PATIENT SIMILARITY NETWORK DRIVEN BY FEDERATED DATA QUALITY AND RESOURCE PROFILING

by

Alramzana Nujum Navaz

Faculty Advisor

Prof. Mohamed Adel Serhani & Dr. Hadeel T. El Kassabi

College of Information Technology

Date & Venue

Friday, 10 Nov 2023

10:00 am -12:00 pm

Building E1 Room: E1-1036

Abstract

Smart and Connected Health (SCH) is revolutionizing healthcare by leveraging extensive healthcare data for precise, personalized medicine. At its core, SCH relies on the concept of patient similarity, which involves the comparative analysis of newly encountered patients with those who exhibit comparable similarities from the existing patient cohort. Yet, this approach faces significant challenges, including data heterogeneity and dimensionality. Our research introduces a multi-dimensional Patient Similarity Network (PSN) Fusion model tailored to handle both static and dynamic features. The static data analysis focuses on extracting contextual information using Bidirectional Encoder Representations from Transformers (BERT), while dynamic features are captured through neural networks and with Long-Short-Term Memory (LSTM) based autoencoders to reduce dimensionality while preserving temporal features. The key to our approach is the novel Similarity Network Fusion (SNF) scheme, that aggregates static and dynamic PSN similarity matrices. Compared to conventional classification methods, our deep learning-based PSN Fusion model demonstrates superior classification accuracy across various patient health outcomes. However, during our evaluation, we identified certain quality issues in the data that need to be addressed at each of the data value chain's processes to maximize the PSN's accuracy. Our Data Quality Management model introduces the data profiling concept to capture, enhance, and validate data quality at every stage of the PSN. We proposed Federated Data Quality Profiling (FDQP), inspired by Federated Learning, to extend the concept of quality profiling to the edge node, ensuring robust data quality assurance in distributed environments. It employs federated feature selection and lightweight profile exchange, to swiftly identify and rectify data discrepancies. Extensive experiments across edge nodes demonstrate the positive impact of FDQP on data quality and the accuracy of Federated Patient Similarity Network (FPSN) models. Finally, we proposed a hybrid resource-aware FPSN solution to effectively combine static and dynamic resource quality profiles with edge reputation data to improve edge node selection. This all-inclusive approach ensures improvements in convergence time, as well as efficient memory, network, and disk usage within FPSN models. In summary, our research integrates the PSN Fusion model, the Federated Data Quality Profiling model, and the Resource-Aware Federated Profiling model to offer a holistic solution. This approach promises transformative benefits by efficiently managing multi-dimensional heterogeneous health data, emphasizing data quality, and optimizing edge computing resources. Ultimately, the goal is to deliver an enhanced healthcare system that empowers healthcare practitioners with efficient and data-driven tools, leading to improved patient outcomes.

Keywords: Smart and Connected Health, Precision medicine, Patient Similarity Network (PSN), Data quality management, Edge Node Selection, Resource optimization, PSN Fusion model, Federated Data Quality Profiling, Resource-Aware Federated Profiling, Federated Learning, Federated Patient Similarity Network (FPSN).