

Health Care Applications with Natural Language Processing



Date:

05 November 2025

Time:

10:00am to 11:00am

Location:

**Core Science Facility,
Memorial University**

Room: CSF-1302

IEEE Newfoundland and Labrador Section Seminar

Dr. Raymond Ng

Canada Research Chair Data Science and Analytics
Director of UBC Data Science Institute
University of British Columbia

Abstract: Unstructured documents often come with embedded structured data. Representing valuable and structured information as tables is popular in health, financial, and many domains. However, manual extraction of structured information from documents typically costs tremendous time and labor, motivating the need for a system for automating the process. After such tables have been extracted, the data can be used for a wide variety of tasks such as question answering and various “down-stream” analytics tasks. In this talk, we will discuss how to leverage ground-breaking pre-trained language models (e.g., BERT, ChatGPT) to develop tools for automated table extraction from various types of documents. We will present different applications from cancer registry reporting, cancer care, and psychiatry hospitalization prediction.

Speaker biography: Raymond Ng is the Canada Research Chair on data science and analytics. He is also the founding Director of the UBC Data Science Institute, and an elected fellow of the Royal Society of Canada. He is currently a co-director of a new UBC AI & Health Network. For both 2022 and 2023, he was named one of the world’s top-75 academic data science leaders by the MIT-based CDO magazine. Ng’s main research area for the past three decades is on data mining, with a specific focus on health informatics text mining, and Natural Language Processing. He has published over 260 peer-reviewed publications on those topics. (H-index 75; total citations 41,000+)