UNIVERSITY of GUELPH

COLLEGE *of* ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

MSc Defence

Wednesday June 12, 2019 at 10AM in Reynolds, Room 1101
Domain Category Information as a Guide for Sentence Ranking to Support Medical Text Summarization

Kasiemobi Maduabunachukwu

Chair: Dr. Gary Grewal
Advisor: Dr. Fei Song
Co-Advisor: Dr. Andrew Hamilton-Wright
Non-Advisory Committee Member: Dr. Rozita Dara

Abstract:

The field of medicine has guidelines which call for medical practitioners to pursue evidence-based medical practice, ones that require them to include the best available evidence from published research in their decision-making process. The exponential growth of electronic biomedical textual resources to read through makes it difficult for medical professionals to keep up with these guidelines. This has led to the recent interest in the development of automatic systems that can generate short, reliable, evidence-based summaries for professionals in the medical practice.

This research contributes to existing research by proposing new strategies to identify the best sentences for summarization based on domain knowledge in a specialized corpus to support evidence-based medical text summarization. The analysis here explored sentence re-ranking guided by document structure measures (PIBOSO type information) and "strength of recommendation" (SORT) categories describing statements of evidence. This work determined that use of these re-ranking strategies leads to different summaries with similar values assessed using the popular ROUGE-L quality metric as those produced by the existing system. To achieve this goal, this thesis also provides automatic grading of SORT evidence expressing sentences using an SVM classifier. A detailed analysis of the performance of both of these contributions is provided, along with future potential applications of the contributions.