MediGraph: A Novel Clinical Assistant from Automated Biomedical Literature Extraction and Knowledge Graphs

Levett, Jordan (School: Vanier College)

Scientific literature is evolving at an accelerated pace. Translating scientific research to medical practice is a challenging task when presented with an overload of publications. Moreover, physicians may not possess the expertise or time to grasp concepts in advanced fields such as artificial intelligence (AI). This study implements machine learning (ML) algorithms and knowledge graphs to develop a literature-based AI clinical assistant. The state-of-the-art Bidirectional Encoder Representations from Transformers (BERT) language model was used to recognize concepts and relationships in unstructured biomedical text. A large biomedical corpus of over 1,500 documents was organized into a training and test set. A knowledge graph was populated with concepts and relationships. Dynamic graph algorithms abstracted meaningful insight via link prediction and neighborhood detection. The system was tested on over 300 annotated documents and achieved a maximum precision and recall of 83.08% and 90.62%, respectively (95% confidence threshold). Annotations and concept relation recognition were performed semi-autonomously. The knowledge graph was dynamically generating newly significant nodes and relationships ($p<0.05$). Furthermore, the viability of this AI assistant can improve standard diagnostic reports through evidence-based medicine. These results suggest that the algorithm can further generalize to unstructured text and can provide informed diagnoses from patient records alongside biomedical literature. The feasibility of the proposed framework allows this web-based system to be readily deployed in the clinical setting or in remote regions where there is a lack of access to medical assistance.

Awards Won:
Fourth Award of $500