Prediction of Venous Thromboembolism using a Hybrid Semantic-based and Machine Learning Approach

The Oakland University and School of Engineering and Computer Science communities are invited to attend Susan Sabra's defense of her Ph.D. dissertation. Seating is limited. RSVP with Katie Loodeen at loodeen@oakland.edu.

Prediction of Venous Thromboembolism using a Hybrid Semantic-based and Machine Learning Approach

Committee: Khalid Malik, Ph.D. (Chair), Mohammad Dadashzadeh, Ph.D., Mohammad Reza Siadat, Ph.D., Eric Fredericks, Ph.D.

Time: **12:00 – 3:00 p.m.**
Date: **Wednesday, June 13, 2018**
Location: **347 EC**

Clinical narratives are a rich source of knowledge as they contain key biomarkers about a patient's health. Predictive analytics for preventive medicine using clinical narratives need accurate semantic and sentiment analysis to extract medical features and classify clinical text using machine learning approaches. This dissertation proposes a new model for diagnosis prediction using a semantic approach and an ensemble classifier. We chose a case study on Venous thromboembolism (VTE) because it is the third most common cardiovascular disorder with a high fatality rate of 25% at first occurrence. Our model matches and maps the concepts of VTE risk factors using Unified Medical Language System (UMLS) medical ontologies and our developed VTE ontology (VTEO). The sentiment assessment module finds adjectives and adverbs that reflect the severity level for each risk factor. To overcome the limitations of using single classifiers, we propose combining the support vector machine (SVM) and Multi-layer perceptron neural network (MPLNN) using an ensemble learning method (stacking). Using a dataset of 250 clinical narratives, the ensemble "MLPNN+SVM" with SESARF achieved a precision of 81.8%, a recall of 81.8%, an F-Measure of 81.8%, and a receiving operating characteristic (ROC) of 80.1% in identifying cases of VTE. These results demonstrate that semantic and sentiment analyses along with ensemble classification of clinical text can help clinicians acquire better evidence from the medical text to make better VTE diagnosis.