LocateMotion: Diabetes Risk Prediction on Heterogenous Electronic Medical Data

In 2020, an estimated 29% of Canadians are living with diabetes or prediabetes. As Diabetes can proliferate into many complications in different physiological systems, creation of a method that tracks the overall health status of those living with diabetes is critical to ensure that patients always receive the care they need. In this study we developed two models on heterogeneous electronic medical record (EMR) data to forecast disease progression and quantify risk levels for diabetic patients. Our first approach quantifies risk scores by stratifying patients with diabetes based on benchmark levels recommended by the American Diabetes Association. Key clinical features are forecasted using ARIMA and recurrent neural network (RNN) models and risk scores are calculated from the forecasted values following clinical guidelines. Our second approach focuses on the progression from diabetes to diabetes-associated complications. In this approach we combine prediction scores from three separately trained RNN models on clinical billing, drug, and lab and exam domains to predict the onset of 10 associated complications such as nephritis and nephropathy. In conjunction, our models provide a comprehensive risk score that is highly interpretable. Ultimately, we aim to integrate our models with online healthcare platforms to provide live analytics to users based on their real-time data. Our models can provide insight to Canadians living with diabetes and their healthcare providers to inform healthy lifestyle decisions and preventative measures for at-risk individuals.