Using ED Chief Complaints to Perform Situational Awareness and Surveillance


Domain Area: Population Health

Problem context: Conventional disease surveillance approaches depend on confirmatory laboratory testing after preliminary diagnosis by a clinician, which may take days of testing and epidemiological analysis before an outbreak is identified and may not be timely enough to provide the information needed to detect and monitor a rapidly evolving problem. In this exercise we will focus on utilizing chief complaints – short strings describing a patient's presenting condition – as a data source for syndromic surveillance. Students will apply text processing techniques to identify cases and will perform time series analysis to look for anomalous events that could indicate an outbreak.

Domain Learning Objective: Identify potential outbreaks using an existing and ubiquitous data source from emergency department admissions.

Data Science Learning Objectives:
1. Ask Questions:
   a. Define syndromes of interest to public health officials
   b. Determine which case definitions could potentially be identified from emergency department (ED) visits
2. Acquire and Assimilate Data
   a. Acquire ED visit data from the University of Utah Emergency Department
   b. Clean, filter, and normalize data
3. Analyze Data and Answer Questions:
   a. Create and apply regular expressions to identify potential cases from chief complaint strings
   b. Characterize signals in the data, including periodicities and increase in utilization
   c. Develop a model for time series analysis that accounts for regular patterns in the data
   d. Validate model using simulated outbreak data supplied by instructors
4. Assess results and Advise:
   a. Discuss limitations of the data source and the models
   b. Determine feasibility of implementing a system for surveillance

Data sets: All ED visits at University of Utah Health from 2000-2016

Data science resources: Jupyterhub notebooks with Python packages such as those described in http://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/