ABSTRACT

Project Title: Can Twitter Catch the Flu

Predicting Influenza Outbreak with Twitter

Project number # D0503

Influenza infections cause 290,000 - 650,000 deaths worldwide each year. Early detection of seasonal flu epidemics may allow more time to mobilize resources and reduce the impact of flu. Currently, government agencies such as Centers for Disease Control and Prevention (CDC) require the collection of Influenza-Like Illness (ILI) data from thousands of healthcare providers, which is slow and expensive. On the other hand, social media is producing massive amounts of data at an unprecedented rate, providing tremendous opportunities for data mining to discover healthcare insights. This study describes a tool for flu surveillance using text mining of Twitter data. Worldwide Twitter data was searched using keyword "flu" during the flu season. The graph of the Flu trend was generated using the daily counts of flu-related tweets, and then compared with CDC Fluview data for correlation to check the validity of current model. After additional development, such social media-based surveillance tools can be useful for the early detection of disease outbreaks and facilitate faster and better medical response preparation.