Unsupervised Text Mining Approach to Early Warning System

Authors : Ichihan Tai, Bill Olson, Paul Blessner

Abstract : Traditional early warning systems that alarm against crisis are generally based on structured or numerical data; therefore, a system that can make predictions based on unstructured textual data, an uncorrelated data source, is a great complement to the traditional early warning systems. The Chicago Board Options Exchange (CBOE) Volatility Index (VIX), commonly referred to as the fear index, measures the cost of insurance against market crash, and spikes in the event of crisis. In this study, news data is consumed for prediction of whether there will be a market-wide crisis by predicting the movement of the fear index, and the historical references to similar events are presented in an unsupervised manner. Topic modeling-based prediction and representation are made based on daily news data between 1990 and 2015 from The Wall Street Journal against VIX index data from CBOE.