Week 9-2: Paper Summaries

CE-510 Seminar: Social Media Mining

Student Name: Jiaqi Guo   NetID: JGR9647

■ Compass: Spatial Temporal Sentiment Analysis of US Election

The objective is to discover sentiment on Twitter towards either the democratic or the republican party at US county and state levels over any arbitrary temporal intervals, using a large collection of geotagged tweets from a period of 6 months leading up to the US Presidential Election in 2016.

And the compass framework includes 5 parts: (1) Tweet Classification Model; (2) Sentiment Analysis Model; (3) Bursty Event Detection Module; (4) Spatial-temporal Analysis Framework; (5) Visualization. For sentiment model they take LSTM to capture long term dependencies.

Possible Improvement Directions:

1. Twitter is still predominant among younger generations. Hence it reflects the voice of the youth generation mostly. It is not possible to get a complete picture until Twitter user becomes popular among all generation.
2. Addressing the multilingual solution on sentiment analysis, expand the analysis to other languages may also be another possible solution.

■ Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion, and sarcasm

The researchers built a deep neural network and then trained it with reams of conversation data from Twitter. The method is also possible thanks to emojis: when people express certain emotions on weibo, they often accompany them with emojis, such as emojis. These emojis, they kind of tag that sentence. From this point forward, the deep learning system can detect that something is wrong with emoticons, even without the knowledge of advanced human irony

The artificial intelligence system, based on deep learning, is called DeepMoji. To train DeepMoji, researchers collected 55 billion tweets, then selected 1.2 billion with 64 common emojis. First, they trained the system to predict which emoji would be used for certain messages, such as happy, sad or happy. The system was then trained to recognize irony. The model classifies expressions as negative, positive and loving, and learns the differences between different emotions. The graph below shows DeepMoji's hierarchical clustering of emojis.

Possible Improvement Directions:

1. In this study, there is a lack of appropriate benchmark data set for emotion analysis, and the emotion data set with the largest number of categories has only 7 emotion categories. To solve this problem, we need to establish a new emotion baseline data set to promote sentiment analysis research.