**Comparing and combining sentiment analysis methods**

The purpose of this paper is to compare 8 different sentimental analysis tools and observe their results for different test sets individually and combined. The sentiment analysis tools used are: (1)Emoticons: detect polarity based on emoticons in the tweet. (2)LIWC: evaluates emotional, cognitive and structural components of a given text based on the use of a dictionary. (3)Sentistrength: a Machine Learning based sentiment analysis tool. (4)Sentiwordnet: obtains scores between [0,1] to evaluate the polarity of a tweet. (5)SenticNet: infers polarity of common sense concepts from NLT at a semantic level. (6)SASA: Machine learning based tool. (7)Happiness Index: calculates the polarity score with respect to the happiness in the tweet. (8)PANAS-t: based on a large dataset of words which associate the tweet with eleven words. The authors made use of two datasets first was a near complete twitter log from 54 million users and 1.7 billion tweets which had topics related to sports, politics, tragedies etc., and second was a manually labeled web 2.0 data. They computed parameters such as accuracy, precision and f-score using the false and true positive scores. The authors then used test data such as plane crash, U.S elections and worldwide epidemic to compute the performance measures and their results showed that for most of the cases Sentiwordnet and senticNet had the highest coverage for every event, which concludes that social media data when analyzed with different tools can be interpreted differently. When the authors tried combining multiple tools the result was only marginally better which states that combination can increase performance but tools must be selected with respect to the data context. They then constructed a webservice know as iFEEl for anyone to test various sentiment analysis methods. However, further research needs to be done on comparing different sentiments and not only positive and negative labels.

**Taste, ties and time: A new social network dataset using Facebook.com**

The purpose of the paper is to study and compare different socio factors of a group of students from Northeastern US. The paper talks about how data was collected from Facebook: (1)They directly downloaded the data from facebook.com and avoided interviewer effects. (2)Data was collected from a complete network which states that the bounded population had a high response rate. (3)Data collected sits on four waves of longitudinal data corresponding to 4 years in college which allows the researchers to observe the change in students network, taste etc. over the time period. (4)Data was collected on multiple social relationships. (5)Facebook friends were collected based on friend list, tie strength and informal reports from Facebook users.(6)Picture friends which tells us additional measures of friendship. (7)Housing information provided by the college. (8)Cultural data for each student such as favorite book, movie, etc. First the authors analyze the social structure of the data for this the authors estimate the income for each student by use of ZIP codes on Facebook, and for students not having zip codes they simply used the university zip code which was not a good way to fill in the missing data. They then found out whether each user was active or no using the last updated activity. Using OLS regression they were then able layout a population demography of factors like gender, race, income etc. and then compared these various factors with each other. They then used the collected cultural data to create a cultural structure for that set of users which included most popular movie, artist etc. The author used these to layout a demography of ties between roommates and dormmates along with picture friends data, sex etc. Further research needs to be done on dealing with network sampling difficulties.