Sentiment Analysis

D. Skrepetos[1]

[1]Department of Computer Science
University of Waterloo

NLP Presenation, 06/17/2015

Outline

Measuring public opinion

- How can one (policymaker, social scientist, etc) **measure public opinion**?
- e.g. what the **outcome** of the **elections** will be?
- e.g. do people **like** the **president**?
- e.g. how do people **feel** about the **economic situation** of the country?

Polls

- Standard solution: Polls
- Many well studied **polling** techniques
- e.g. telephone polls
- **Time-consuming** and **money-demanding** even for calling 1,000 people
- **Limits** on the number of **questions** and on the number of **people**

Text-based social media

- Alternate solution: **text-based social media**
- e.g. facebook, twitter, blogs
- **Millions** of people post their **thoughts** and **opinions** on many **topics**
- Easy to **access**
- Easy to **analyze**
- Can ask a lot of **questions**
- Can ask a lot of **people**

Twitter

- **Millions** of people post tweets in **Twitter** that **reveal** their **opinions** on **public matters**
- Tweets are **small**
- **Sentiment analysis techniques** work well, especially on **small** sentences
- Use sentiment analysis on tweets to **measure public opinion**!
- **Estimate** accuracy by **comparing** the measures with public opinion polls

Twitter data

- 1 billion tweets between 2008 and 2009 **collected**, mostly from U.S. and in English
- Twitter API and Gardnhose real-time stream used for **collection**
- 100,000 to 7 million tweets **per day**
- Twitter **grew** 50 times over the period

Twitter data problems

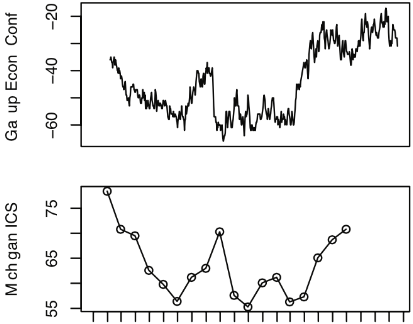
- Does not filter out non-U.S. users and non-English tweets
- Twitter **demographics** do not necessarily represent **general population**
- So, better **preprocessing** needed

Consumer confidence

- Consumer confidence: how **confident** people are for **country** and **household finances**
- Related to **economic activity**
- Needed for **economic policy making** and **business planning**

Consumer confidence poll data

- Telephone **surveys**
- Index of Consumer Sentiment (ICS) extensively studied
- Five questions monthly to several hundred people nation-wide
- Gallup's Economic Confidence index
- Two questions daily

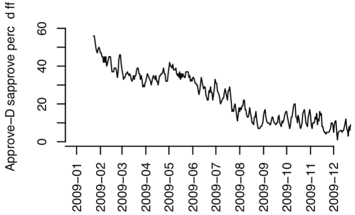


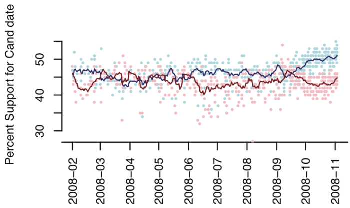
Political opinion

- Political opinion: **whom** the voters will **vote** *and* how they **assess** a **politician's job**
- **Presidential** job approval
- Winner of 2008 U.S. **elections**

Political opinion poll data

- Telephone **surveys**
- Gallup's daily tracking poll for **presidential job approval** of Barack Obama as 3-day rolling averages
- Tracking polls during 2008 U.S. **presidential election cycle** asking voters if they will vote Barack Obama or John McCain
- 46 different polls from Pollster.com



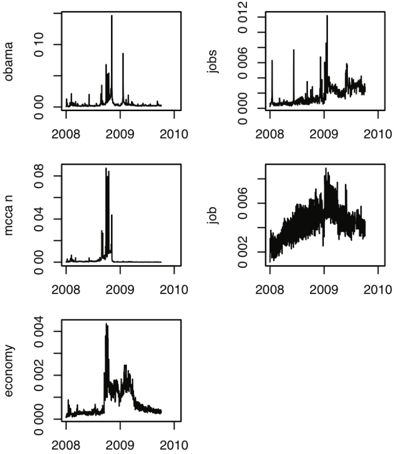


Method outline

- **Message retrieval**: find the relevant to the topic tweets
- **Opinion estimation**: decide if the tweets are positive or negative

Message retrieval

- **Manually** determine topic **keywords**
- e.g. for presidential approval: *Obama*, *Barack*, *president*
- Retrieve all tweets containing those **keywords**



Opinion estimation

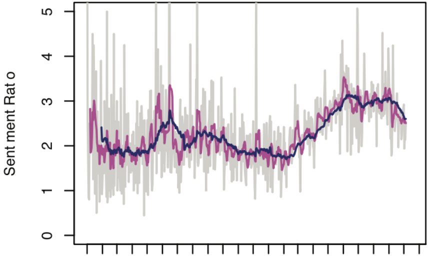
- If a tweet contains a positive word, it is positive
- Similarly for negative
- Find positive and negative words using the OpinionFinder's lexicon
- Therefore, a tweet can be both positive and negative
- Simply **count** number of positive and of negative tweets
- Measure **sentiment score** x_t on day t as the ratio of positive vs negative tweets
- e.g. if in day $t = 2$ there are 5 positive and 4 negative tweets, then $x_2 = \frac{5}{4} = 1.25$

Problems with opinion estimation

- OpinionFinder's lexicon assumes **part-of-speech tagger**, but **none** is used
- e.g. **noun** *will* counts as positive, but it is mostly used as **verb**
- "The global economy *will* be destroyed" is counted as a positive tweet, but it is negative
- Also, OpinionFinder's lexicon assumes **well-written English**, but tweets use **heavily informal English** and **emoticons**
- e.g. "Barack Obama ☺" or "Barack Obama is THE MAN!"

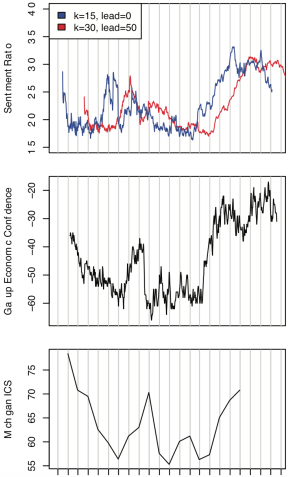
Moving average aggregate sentiment

- The **sentiment ratio** of tweets is **more volatile** than polls'
- That's not a problem, but we have to **compare** it with polls
- So, let's **smooth** it
- That is, we take the **average** over the last k days, MA_t
- $MA_t = \frac{1}{k}(x_{t-k+1} + x_{t-k+2} + +x_t)$
- e.g if k=2, t=5, $x_4 = 2$ and $x_5 = 1$, then $MA_5 = \frac{1}{2}(2+1) = \frac{3}{2} = 1.5$



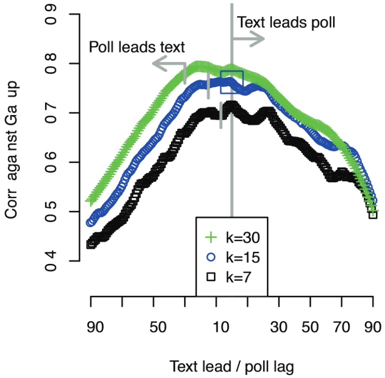
Results

- **Sentiment ratio** captures the **trends** in the data
- For example for the keyword *jobs* and the economic confidence indexes



Forecast

- Can the **sentiment ratio** predict **future trends** before polls?
- **Shift** the x_t sequence L times to the left
- Measure **correlation** of the average and the polls



Conclusion

- The method can successfully **predict public opinion** by inspecting tweets
- The **sentiment analysis techniques** are **simple**
- So, there may be room for **improvement** by using more **advanced** techniques
- Better suited **lexicon**
- Why tweets predict public opinion?
- Twitter continually **evolves**

Outline

Predicting stock market

- Predict **stock market values**
- Early research on **random walk theory** and the **Efficient Market Hypothesis**
- EMH says that markets are driven by **news**, so they are **unpredictable**
- Socionomic Theory of Finance, behavioral economics, and behavioral finance **examined** EFH
- Later, stock market values do not follow random walks, so the **can** be **predicted**
- Even later, online social media predict changes in various **economic** indicators

Emotions and stock market

- **Emotions** play important role in human **decision-making**
- Behavioral finance gave **proof** that emotion and mood affect financial decisions
- Need **large** and **early** public **mood** surveys
- Large **surveys** of public mood are **expensive** and **time consuming**
- As in previous paper, use tweets to **predict** the **stock market**

Data

- 10 million tweets collected
- **Preprocess** tweets to **remove** stop-words, punctuation, etc and **filter out** tweets without **feeling** words
- **Compare** against daily DJIA (Dow Jones Industrial Average)

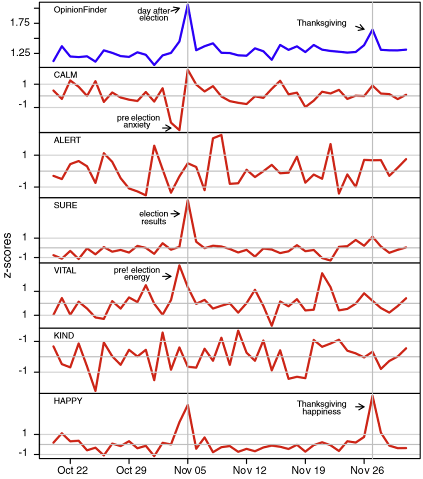
Mood assessment tools

- First step: Use **mood assessment tools** to build a **mood** time series vector
- OpinionFinder returns positive vs negative **mood**
- GPOMS returns measures of **6** different **mood aspects**
- So, we have a **time series vector** with **7 mood aspects** per day

OpinionFinder and GPOMS

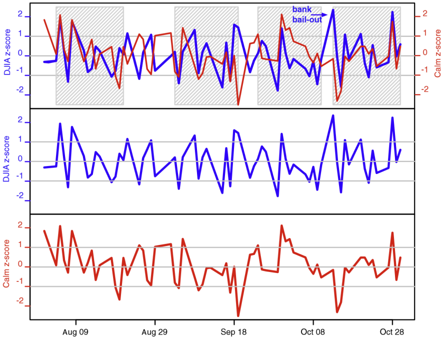
- OpinionFinder decides if a sentence is positive or negative
- However, it **ignores** the **richness** and the **multi-dimensionality** of human mood
- GPOMS returns **6 mood aspects**: Calm, alert, sure, vital, kind, and happy

Evaluation of the mood predictor for social events



Granger causality analysis

- Second step: **Evaluate** how well the **mood predictor** *predicts* DJIA
- Use the **Granger causality analysis**
- If X **causes** Y, then **changes** in X occur **before** changes in Y



- So, **calm** *can* **predict** DJIA

Self-organizing fuzzy neural network

- Granger analysis is **linear**, but **relation** between **public mood** and **stock market** values probably is **not linear**
- Granger analysis says that DJIA is related **only** with calm, but DJIA could be dependent on happy **combined** with calm
- Use **self-organizing fuzzy neural network** to *discover* **non linear** relations
- **High accuracy** and **interpretability**
- A **combination** of **happy** and **calm** *predicts* DJIA

Conclusion

- **Generalization** for other **countries** and **languages**
- Why the **correlation** between tweets and DJIA **exist**?
- How to avoid **manipulation**?

Outline

1. From Tweet to Polls: Linking Text Sentiment to Public Opinion Time Series
2. Twitter Mood Predicts the Stock Market
3. Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach

Predicting stock market

- **Huge** number of **reviews**, ratings and recommendations on **blogs** and **social networks**
- Given a **review**, automatically **label** it as positive or negative
- Important for **businesses** for selling products and discovering opportunities
- **Numerous reviews** for a lot of **different** *topics*
- e.g. movies, cellphones, guitars

Multiple classifiers

- Build a **classifier** for the **polarity** of a review of **any** topic
- The **polarity words** in a review are **different** across **topics**
- e.g. reviews for kitchen appliances: *malfunctioning*, *reliable*
- e.g. reviews for DVDs: *thrilling*, *horrific*
- So, data **distributions** are **different** across **domains**
- **Difficult** to gather **annotated training data** for **all** of topics
- Does not exploit **information shared** by **different** domains

One classifier

- We would like to build and train a classifier for **only one** topic
- **Train** it for **one** topic s.t. it **generalizes** well on others
- *Discover* intermediate **abstractions** that exist across **all** the domains
- In reviews **words** will be **different**
- However, the **words** are used for **abstract concepts** like *quality*, *price*, *costumer service*
- **Domain adaptation**

Domain adaptation

- Build **classifier** with training set from p.d. S that **generalizes** well to a test set from p.d. T with $p_S \neq p_T$
- Algorithms for discovering **intermediate representations** in a **hierarchical** manner
- **Unsupervised** learning for building each **level** of **features** based on the **previous** level
- **Huge** number of **unlabeled reviews**

Amazon data

- 340,000 **reviews** across 22 topics
- Vast **disparity** in the number of reviews **per topic**
- **Heterogeneous**, **heavily unbalanced**, **large-scale**
- **Comparison** against 3 well-studied **methods**
- The **new** method is **better** than the **previous** 3

Method

- First step: *extract* **higher-level features** with **unsupervised** learning from the reviews of **all** domains using a Stacked Denoising Autoencoder (SDA)
- SDA works in an **unsupervised** way, so it can use **all** the domains
- SDA gives **non-linear** mapping, capturing **complex** relations
- SDA gives **sparse** representations
- Second step: build a **linear classifier** using Support Vector Machines (SVM)
- SVMs work **well** on **sentiment analysis**

Conclusion

- First **paper** to use **domain adaptation** techniques in **sentiment analysis**
- Very **simple** methods SDA and SVM used as **black boxes**
- Can other **feature extraction** methods be applied?

References I

B. O'Connor, R. Balasubramanyan, B.R. Routledge, and N.A. Smith.
From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series.
In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 122-129, 2010.

J. Bollen, H. Mao, X. Zeng.
Twitter mood predicts the stock market.
Journal of Computational Science, 2(1): 1-8, 2011.

X. Glorot, A. Bordes, and Y. Bengio.
Domain adaptation for large-scale sentiment classification: A deep learning approach.
In Proceedings of the 28th International Conference on Machine Learning , 513-520, 2011.