Sentiment Analysis: a case study
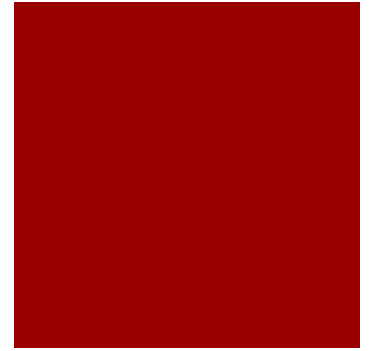
Giuseppe Castellucci
castellucci@ing.uniroma2.it

Web Mining & Retrieval a.a. 2013/2014

Outline

- Sentiment Analysis overview

- Brand Reputation

- Sentiment Analysis in Twitter

- Modeling SA in Twitter

Sentiment Analysis

- **Opinion Mining** or also **Sentiment Analysis** is the computational study of opinions, sentiments and emotions expressed in texts

- It deals with *rational* models of *emotions* and *trends* within user communities

- It is the detection of **attitudes**

- Why opinion mining now?
  - Mainly because of the Web
  - Huge volumes of opinionated text

Sentiment Analysis

- Opinion Mining or Sentiment Analysis involve more than one linguistic task
  - **An opinion is a quintuple**

- What is the opinion of a text
  - Who is author (or **opinion holder**)
  - What is the opinion **target** (Object)
  - What are the **features** of the Object
  - What is the **subjective position** of the user
  - **When** the opinion is expressed

- **Importance of opinions**
  - Opinions are important because whenever we need to make a decision, we want to hear others' opinions

Sentiment Analysis

- Does the opinion of an user matter?

- One can express personal experiences and opinions on almost anything, at review sites, forums, discussion groups, blogs

- Authority and reputation of users are key factors to understand and account for their opinions

- Mining opinions expressed in the user-generated content
  - A very challenging problem
  - Practically very useful

Sentiment Analysis – applications

- **Businesses and organizations**: product and service benchmarking. Market intelligence
  - Business spends a huge amount of money to find consumer sentiments and opinions
    - Consultants, surveys and focused groups

- **Individuals**: interested in other's opinions when
  - Purchasing a product or using a service
  - Finding opinions on political topics

- **Ads placements**: Placing ads in the user-generated content
  - Place an ad when one praises a product
  - Place an ad from a competitor if one criticizes a product

- **Opinion retrieval/search**: providing general search for opinions
  - Search engines do not search for opinions
  - Opinions are hard to express with a few keywords

An Example

- *"I bought an iPhone a few days ago. It was such a nice phone. The touch screen was really cool. The voice quality was clear too. Although the battery life was not long, that is ok for me. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, and wanted me to return it to the shop. …"*

- **What do we see?**
  - **Opinions**, **targets of opinions**, **and opinion holders**
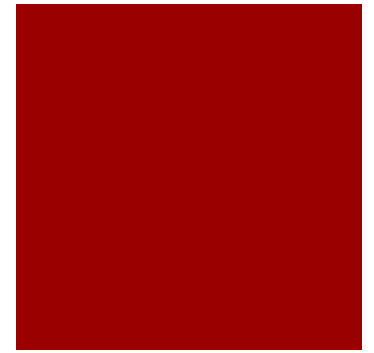
Brand reputation

- Monitoring (online) the reputation of one or more brands
  - An active part today of **brand management**

- Brand reputation is
  - A **continuous** and **real-time** process (potentially)

- Brand reputation allows to find
  - **What** people think on brands
  - **Where** people expresses these opinions
  - What is the **sentiment** expresses

Who is interested on?

- Mainly companies

- Imagine it is possible to know what people think on your brands

- Automatically looking at the web

- In order to adjust company strategies

- In order to better satisfy user needs

- … in order to increase turnover!

The role of SA

- What can do SA for Brand Reputation?
  - In BR opinions are a key component
  - SA mine exactly opinions and sentiment

- A key component for BR is Sentiment Analysis at different levels
  - **Document**: produce general ideas on what people think
  - **Sentence**: more detailed analysis

- **Topics**: interested only in some specific topics during time

The role of the web

- The web cannot be controlled (hopefully)

- It is impossible to avoid a negative news is propagated if it is relevant to someone

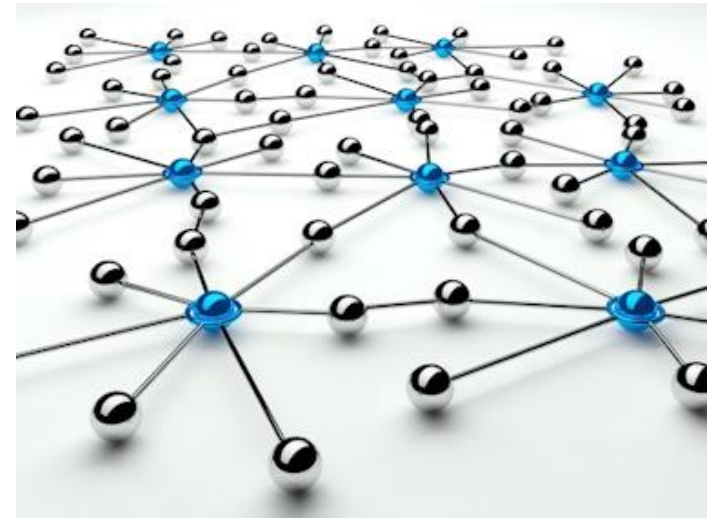- The web can only be **monitored**

- That is, people can be **monitored**
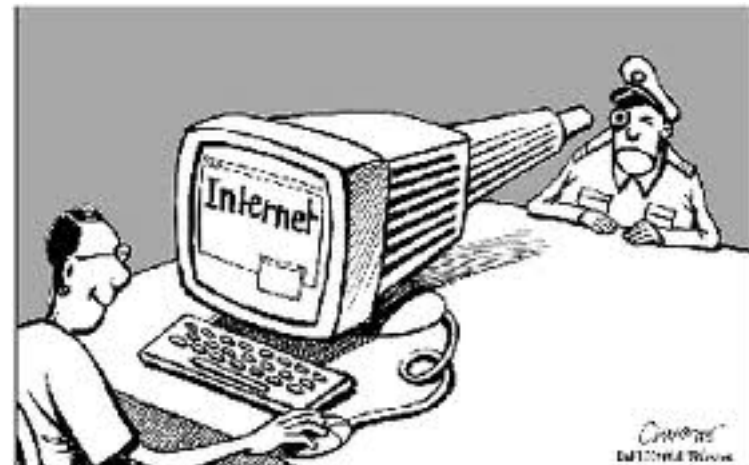
Web 2.0 and Social Media

- **Word-of-mouth on the Web**
  - User-generated media: One can express opinions on anything in reviews, forums, discussion groups, blogs

- Opinions of global scale, no longer limited to
  - Individuals: one's circle of friends
 - Businesses: small scale surveys, tiny focus groups, etc.

Twitter

- Twitter is an online social networking and micro-blogging service

- Who is on Twitter?
    - Common people
    - VIPs
    - Politicians
    - Companies

- Sharing of short (140 chars) messages
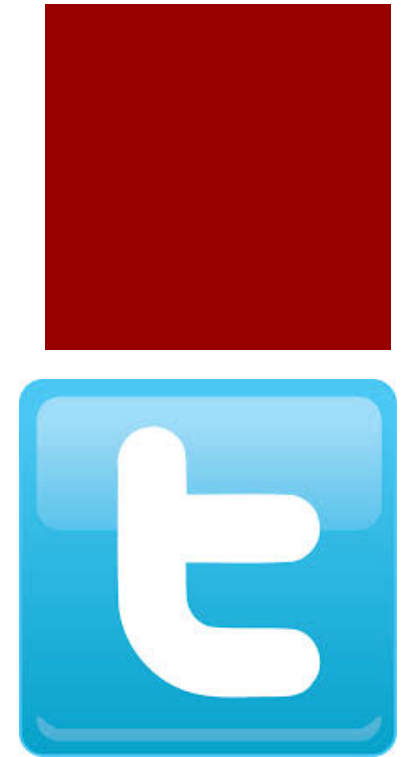
Twitter numbers

- Active users
  - 645,750,000

- Average number of tweets per day
  - 58 million

- Number of active Twitter users every month
  - 115 million

- Number of tweets that happen every second
  - 9,100
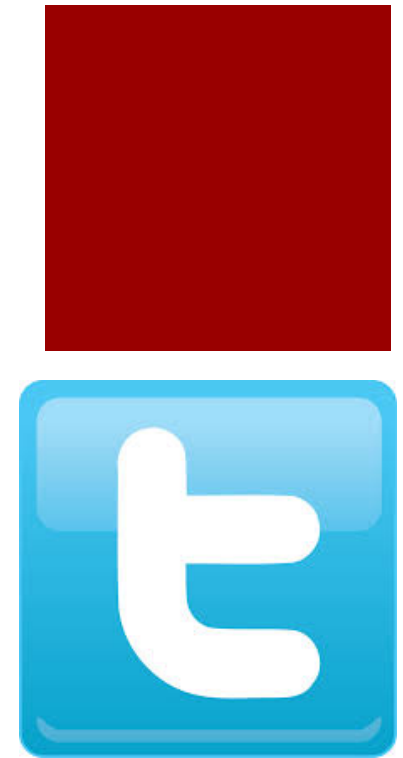
Examples of tweet

Shannyn Moore @shannynmoore · Jul 25
OK, I very rarely see films. I watched "Ip Man" about Yip Man - **Wing Chun** master. I loved it. See it. **Wonderful**.

Collapse   ↩ Reply   ♺ Retweet   ★ Favorite   ••• More

RETWEET
1

Ahmad Alnassar @A_Alnassar_A · May 23
Wonderful moments with friends at ferrari world 👍 @ Ferrari World
instagram.com/p/oWKeheKVNr/

Collapse

6:37 PM - 23 May 201

David Lebovitz @davidlebovitz · May 6
Back at **Apple** Store for my phone : (-- hope the third time is a charm, and worth missing lunch over

Collapse   ↩ Reply   ♺ Retweet   ★ Favorite   ••• More

FAVORITES
3

9:49 PM - 6 May 2014 · Details

SA in Twitter

- Capture the **sentiment** expressed in short messages

- No limit to the range of information conveyed by tweets
  - Often used to share opinions and sentiments

- Twitter is seen today as an instrument to spread messages

- VIPs use it to communicate with fans
  - They want to know if they are appreciated or not

- Politicians use it to make their claims
  - They want to know the reaction of people to them

- Companies use it to create a direct channel with customers
 - They want to have an indicator of user's happiness

SA in Twitter: difficulties

- **Difficulties** due to
  - Shortness of messages, really poor *signal*
  - Usage of informal language
  - Frequent misspellings
  - Use of emoticons
  - Use of hashtags and mentions to users

- Moreover,
  - Sentiment depends on the opinion holder
  - It depends also on the social context
  - E.g. during elections people express messages with a political topic

Case study
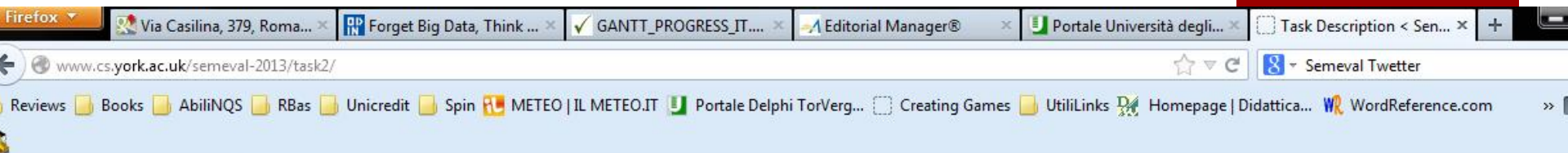
Home   Data   Contact

Sentiment Analysis in Twitter

Task Description

SemEval-2013: Sentiment Analysis in Twitter
- **Please register: http://www.cs.york.ac.uk/semeval-2013/index.php?id=registration**
- **Join the Google group: semevaltweet-2013@googlegroups.com**
- **Full training, development and test datasets, and a scorer/format checker, have been released.**
- **Having problems downloading the data? Try the latest version of the Python script.**

Abstract:

In the past decade, new forms of communication, such as microblogging and text messaging have emerged and become ubiquitous. While there is no limit to the range of information conveyed by tweets and texts, often these short messages are used to share opinions and sentiments that people have about what is going on in the world around them. We propose this task and the development of a twitter sentiment corpus to promote research that will lead to a better understanding of how sentiment is conveyed in tweets and texts. There will be two sub-tasks: an expression-level task and a message-level task; participants may choose to participate in either or both tasks.

- Task A: Contextual Polarity Disambiguation

### ⊠ Contact Info

Organizers

Theresa Wilson (taw@jhu.edu)
Zornitsa Kozareva (kozareva@isi.edu)
Preslav Nakov
(nakov@comp.nus.edu.sg)
Alan Ritter
(aritter@cs.washington.edu)
Sara Rosenthal
(sara@cs.columbia.edu)
Veselin Stoyanov (ves@cs.jhu.edu)

⊠ Other Info

Announcements

What is SemEval

- SemEval
  - Evaluation Campaign of the Association for Computational Linguistics

- Evaluations are intended to explore the **nature of meaning** in language
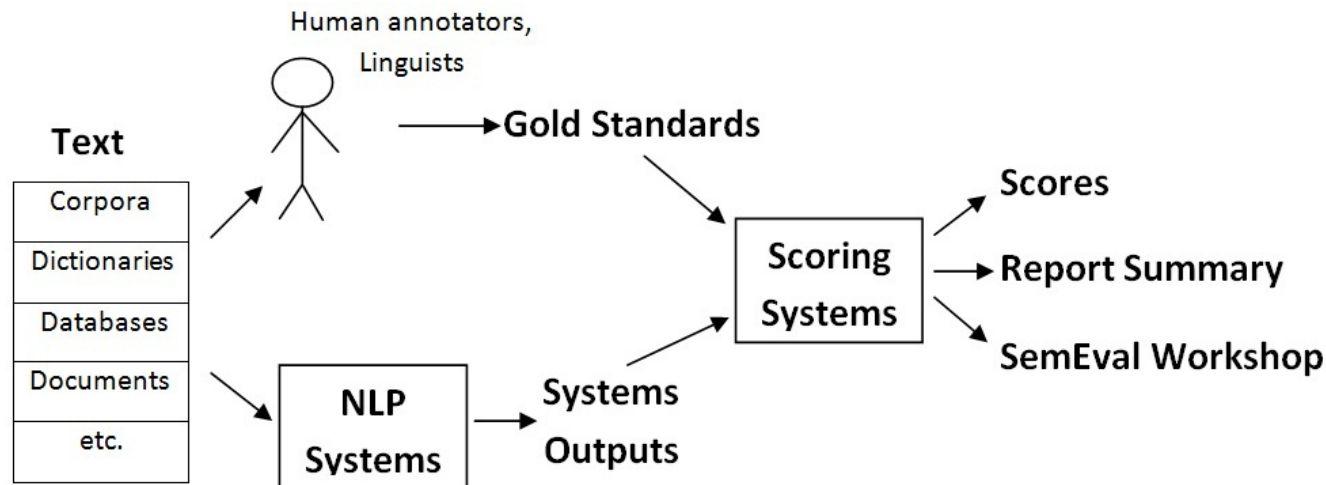
What is a task?

- A task is related to a natural language problem

- It is intended to let researchers study and discuss

- To compare and share solutions in computational approaches to NLP

1. Evaluation of Compositional Distributional Semantic Models on Full Sentences through Semantic Relatedness and Entailment
2. Grammar Induction for Spoken Dialogue Systems
3. Cross-Level Semantic Similarity
4. Aspect Based Sentiment Analysis
5. L2 Writing Assistant
6. Supervised Semantic Parsing of Spatial Robot Commands
7. Analysis of Clinical Text
8. Broad-Coverage Semantic Dependency Parsing
9. Sentiment Analysis in Twitter
10. Multilingual Semantic Textual Similarity
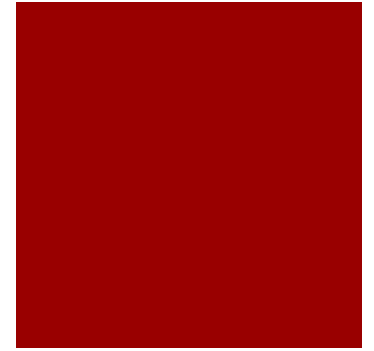
Who cares? Why?

- All of us!

- Because it is an useful way to share
  - Ideas
  - Techniques
  - Compare systems

- Systems in SemEval can be of interest for companies

Sentiment Analysis in Twitter

- Task A

  - **Contextual Polarity Disambiguation**

    - Given a tweet and a span on it classify the instance with respect to *positive*, *negative* o *neutral*

- Task B

  - **Message Polarity Classification**

    - Given a tweet classify it with respect to *positive*, *negative* o *neutral*

Task A examples

- @_Ms_R have you watched Four Lions? Saw it again last night, love it

- its not that I'm a GSP fan, i just hate Nick Diaz. can't wait for february.

- Going to Singapore tonight :) Excited for Skyfall + penny boarding tomorrow!

- Cowboys will beat the falcons sunday #iStamp

- Contraband was prolly the worst below average movie I ever sat and finished watching in my life

Task B examples

- **negative** I officialy hate Windows 7, it just sucks on my laptop. Back to XP tomorrow! (and also a huge delay for the movie again :/, was so close!)

- **negative** Desperation Day (February 13th) the most well known day in all mens life.

- **positive** #NBA I'm excited :)

- **positive** Eli manning best 4th quarter QB in the game!

- **neutral** Harry Redknapp is being heavily linked with the position as Blackburn manager today. However, QPR may have begun courting him already...

- **neutral** 20 June 2012, out European Tour party drove the fabulous road from Davos to Stelvio. This is the last little bit... http://fb.me/1ZkWMVySv

Let's do it!

- Objective: find a possible solution to task B

- A little recipe
  - First of all, *think* on the problem
  - *Think* on the available information
  - *Think* on how to model the problem!
  - Choose the algorithm
  - Experiment, experiment and... experiment

A lot of thinking before experimenting
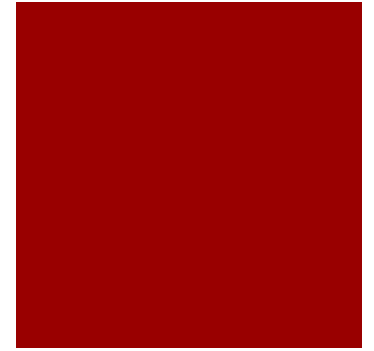
Think on the problem

- **Problem definition**:
  - Classify a tweet w.r.t. *positive*, *negative* and *neutral* classes

- What is a tweet?

- What does make a tweet different from other texts?

Think on the available information

- 10205 training tweets and 3813 testing tweets
  - Three classes: positive, negative and neutral

- Do you have some external resource?

- Do you have a NL processor?

Resources

- SentiWordnet
  - lexical resource for opinion mining
  - assigns to each synset three sentiment scores
    - Positivity
    - Negativity
    - Neutrality

SentiWordNet — good — Search!

ADJECTIVE

good#1   01123148

having desirable or positive qualities especially those suitable for a thing specified; "good news from the hospital"; "a good report card"; "when she was good she was very very good"; "a good knife is one good for cutting"; "this stump will make a good picnic table"; "a good check"; "a good joke"; "a good exterior paint"; "a good secretary"; "a good dress for the office"

P: 0.75 O: 0.25 N: 0   Feedback on SentiWordNet values: [They are OK.] [Suggest your values.]

good#2   full#6   00106020

having the normally expected amount; "gives full measure"; "gives good measure"; "a good mile from here"

P: 0 O: 1 N: 0   Feedback on SentiWordNet values: [They are OK.] [Suggest your values.]

Resources

- MPQA Subjectivity Cues Lexicon

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proc. of HLT-EMNLP-2005.

Riloff and Wiebe (2003). Learning extraction patterns for subjective expressions. EMNLP-2003.

- Home page: http://www.cs.pitt.edu/mpqa/subj_lexicon.html

- 6885 words from 8221 lemmas
  - 2718 positive
  - 4912 negative

- Each word annotated for intensity (strong, weak)

Resources

- Bing Liu Opinion Lexicon

Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. ACM SIGKDD-2004.

- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar
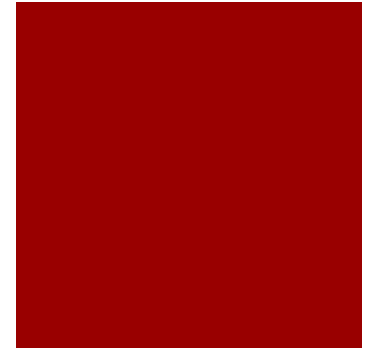
- 6786 words
  - 2006 positive
  - 4783 negative

NL processor

- E.g. tokenizers, part-of-speech taggers, parsers

- Tweets often
  - do not follow common language rules
  - convey their message through emoticons
  - …or hashtags

- Need for specialized processor

Ad-hoc NL processor

- Normalize emoticons, links
  - #NBA I'm excited SMILE_HAPPY
  - 20 June 2012, out European Tour party drove the fabulous road from Davos to Stelvio. This is the last little bit... LINK

- POS tags are useful for Twitter SA
  - You can know that a word is an adjective (e.g. good, bad) a verb (like, dislike, hate), a noun, etc…

- Parsers are less useful
  - A tweet hardly have a syntactic structure

A Twitter NLP chain

- A simple NLP chain for Twitter

- Apply a Preprocessing stage

- Apply a standard Natural Language Processor

- In our case,
  - Tokenizer (correctly handles hashtags)
  - POS-tagger

Tweet

Preprocessing

Tokenizer

POS-Tagger

Analyzed tweet
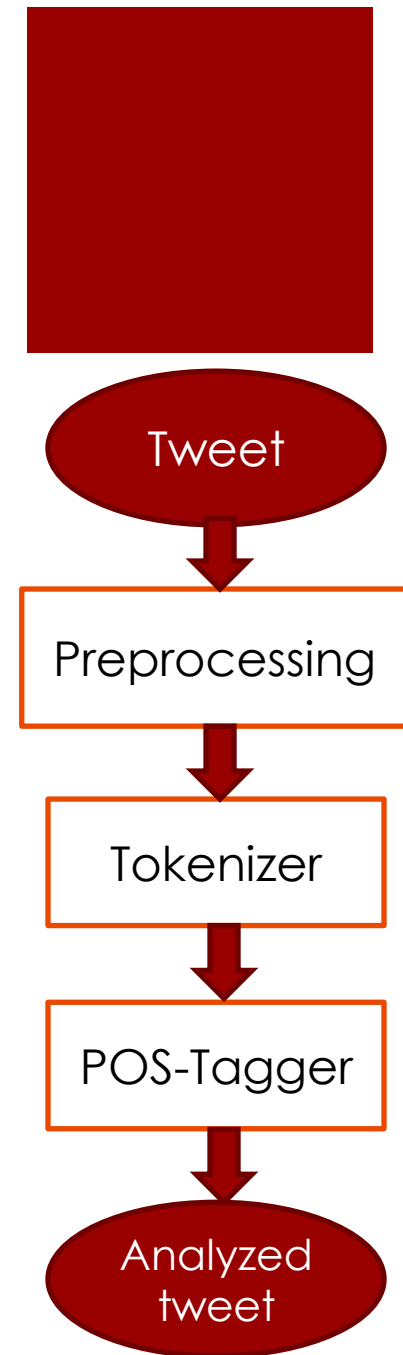
Think on the model

- We want to apply learning methods

- Thus, capture similarity between examples

- How can we capture similarity of tweets with available information?

- Can we apply kernel methods?

Bag-Of-Words

- Emphasize pure lexical information

- Word overlap between tweets
  - It measures how two instances are similar by their sharing of words

- Boolean weighting schemas
  - A word is present or not

- To capture the role of adjectives, nouns, etc…
  - each dimension represents a <lemma,POS> pair

Distributional features

- How to generalize lexical information?
  - Construct a Wordspace

- Download a lot of tweets (3 million)

- Unsupervised analysis through LSA
  - Construct a word-by-context matrix
  - Apply SVD
  - Approximate to k=250 dimensions

- Projects words of a tweet in the Wordspace

- Construct a vector for an example by the sum of the vectors of the words composing it
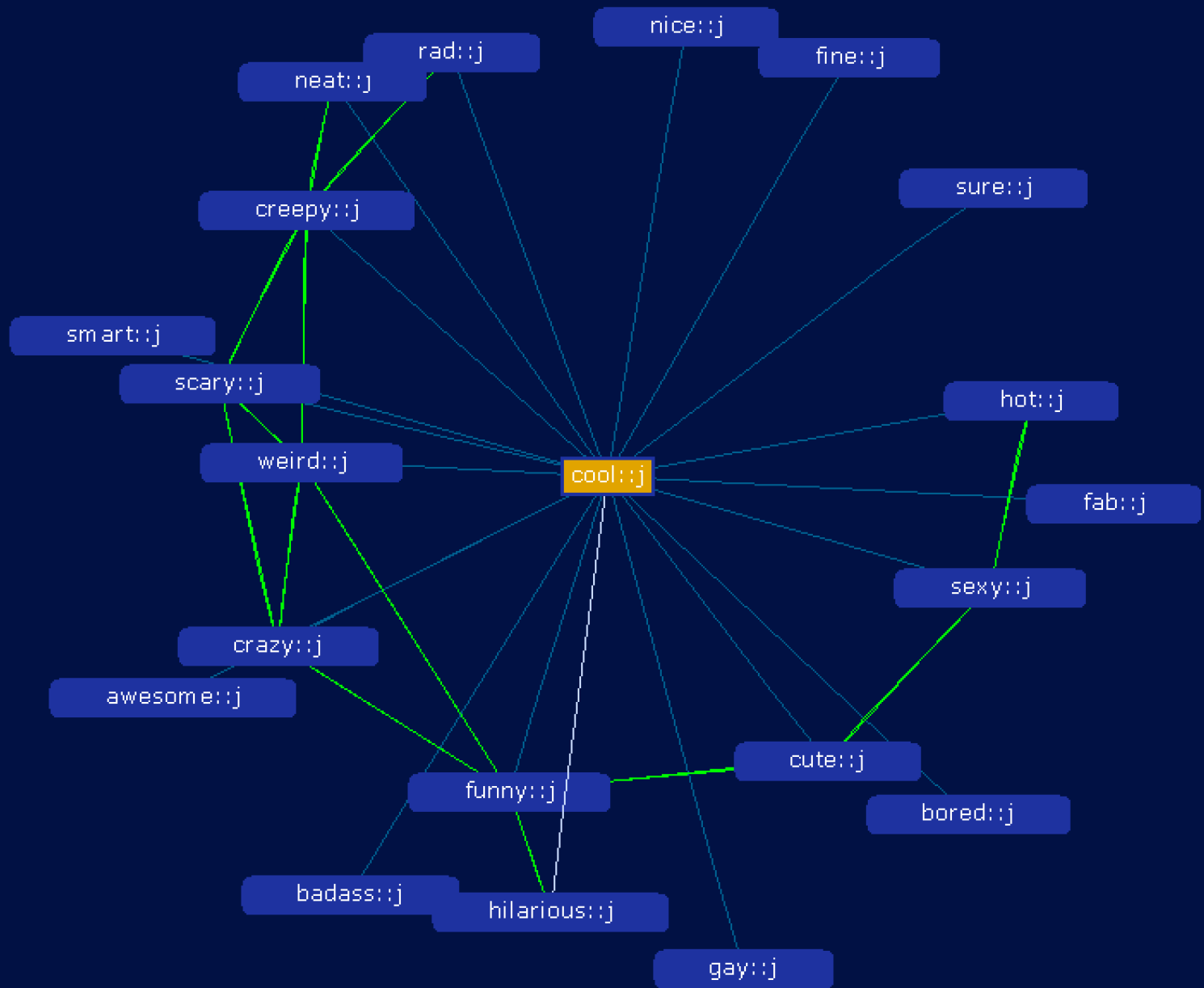
Kernel functions

- Linear kernel
  - Estimates similarity as the dot product between vectors
  - We apply it to the BOW vector

- RBF kernel
  - Estimates similarity as $\quad K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{||\mathbf{x} - \mathbf{x}'||_2^2}{2\sigma^2}\right)$
  - Apply it to LSA vector

- Combine kernel contributions
  - Linear combination of kernels is a valid kernel
  - Our final Kernel is the sum of the linear kernel and of the RBF kernel

Choose the algorithm

- Choose a proper learning algorithm

- We have to select a kernel based algorithm

- Support Vector Machines
    - Three classes, what strategy adopt?
    - Here, **One Vs. All**
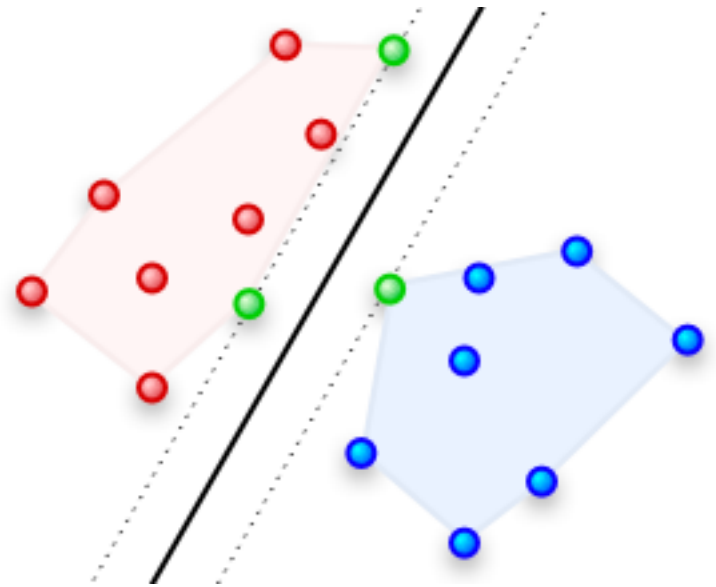
Experiments

- Tune parameters
  - E.g. C for SVM

- Choose a strategy for tuning
  - Fixed split
  - N-fold

- Try other ideas
  - Other kernels, e.g. polynomial kernel on BOW vector
  - New features, e.g. we do not manage negation

- Choose the best combination on the validation set and

- Test on unseen data
  - Compute final performance measure

SAG results in task B

- Performance measure is the mean of the F1-Measure for the positive and negative classes

- A Support Vector Machine implementation with One Vs All strategy achieve **0.65** F1

- The best system at SemEval 2013 **0.69** F1
  - They used a lot of extra information
  - E.g. manual coded resources regarding sentiment of words

Task A

- Now, it's your time!
  - **Strongly suggested** even if not mandatory

- Apply the same recipe to task A

- Spend time in *thinking*, *thinking* and again *thinking* on what you are doing

- Then, if you want to do some experiment ask us for data

References

- http://aclweb.org/aclwiki/index.php?title=SemEval_Portal

- UNITOR: Combining Syntactic and Semantic Kernels for Twitter Sentiment Analysis. Castellucci et Al., 2013, SemEval 2013

- Robust Language Learning via Efficient Budgeted Online Algorithms. Filice et Al, 2013, SENTIRE (ICDM) 2013

- Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Bing Liu, Springer 2011

- SemEval-2013 Task 2: Sentiment Analysis in Twitter. Nakov et Al, 2013, SemEval 2013

- Sentiment analysis of Twitter data. Agarwal et Al. 2011, LSM 2011

- Twitter as a corpus for sentiment analysis and opinion mining. Pak, Paroubek 2010, LREC 2010

- Robust sentiment detection on twitter from biased and noisy data. Barbosa, Feng 2010, Coling 2010