Sentiment Analysis: modern methods, resources and applications

Davide Buscaldi,
LIPN - Université Paris 13

davide.buscaldi@lipn.univ-paris13.fr

Cagliari
16/05/2017 - 23/05/2017

Plan - day 1 (today)

- Computational Semantics: an introduction
- From senses to emotions: Sentiment Analysis
  - Sentiment Lexicons
  - A Sentiment Analysis system: some insights of IRADABE
  - Example application: natural disaster management
- Afternoon: experiences with datasets from previous challenges (Evalita, SemEval)

Plan - day 2

- Fine-grained sentiment analysis
- Recent advances:
  - Deep learning and Sentiment Analysis
  - Machine reading and Sentiment Analysis
  - Lived Experiences
- Detecting fake reviews
- Afternoon: discussion on day 1 experiences and final test

Computational Semantics

- Objective: to *understand* a sequence of symbols
  - 习近平：这次访英将两国关系提高到新水

Xi Jinping

Understand (according to Frege) = link symbols (*references*) to concepts (*senses*)

=> Need Dictionaries

WIKI LOVES AFRICA

Wiki Loves Africa: share African cultural fashion and adornment pictures with the world!

Xi Jinping

*Este es un nombre chino; el apellido es **Xi**.*

Xi Jinping (chino tradicional: 習近平, chino simplificado: 习近平, pinyin: *Xí Jìnpíng*, AFI: [ɕǐ tɕînpʰǐŋ]), (Pekín, 15 de junio de 1953) es el secretario general del Comité Central del Partido Comunista de China, Presidente de la Comisión Militar Central,[1] y, desde el día 14 de marzo de 2013, Presidente de la República Popular China.[2]

Xi Jinping
习近平

Xi Jinping

Secretario General del Partido Comunista de China	
Actualmente en el cargo	
Desde el 15 de noviembre de 2012	
Predecesor	Hu Jintao
Presidente de la República Popular China	
Actualmente en el cargo	
Desde el 14 de marzo de 2013	
Primer ministro	Li Keqiang
Vicepresidente	Li Yuanchao
Predecesor	Hu Jintao

Índice [ocultar]

1 Biografía
2 Carrera política
3 Véase también
4 Referencias
5 Enlaces externos

Biografía [editar]

Xi, pertenece a la etnia han, y es hijo de Xi Zhongxun, ex viceprimer ministro de China y uno de los fundadores de la guerrilla comunista que operó en el norte de China contra el Kuomintang. Xi nació en Pekín en 1953. Durante la Revolución Cultural fue enviado a trabajar al campo. Después estudiaría ingeniería química en la Universidad de Tsinghua, en Pekín, donde también obtuvo el doctorado en teoría marxista en la Escuela de Humanidades y Ciencias Sociales (Universidad Tsinghua).[2]

Carrera política [editar]

En 1974 se unió al Partido Comunista de China. En su carrera política ha recorrido algunas de las provincias costeras con mayor desarrollo. En 1999 fue nombrado gobernador de la provincia de Fujian. En el año 2000 fue nombrado secretario del PCCh de la provincia de Zhejiang.[2]

Después de ganarse reputación de luchador contra la corrupción, en marzo de 2007 fue nombrado secretario del PCCh en Shanghái, reemplazando a Chen Liangyu, destituido del cargo acusado de corrupción.

En octubre de 2007 fue nombrado miembro (el sexto, según el protocolo), del Comité Permanente del Buró Político del Comité Central del Partido Comunista de China. El 11 de febrero de 2008 fue nombrado encargado de los Juegos Olímpicos de Pekín, confirmando de manera extraoficial que Xi sería nombrado próximo vicepresidente.

El 15 de marzo de 2008 fue elegido Vicepresidente de la República Popular China, en la XI Asamblea Popular Nacional de China, al mismo tiempo que Hu Jintao fue reelecto presidente.

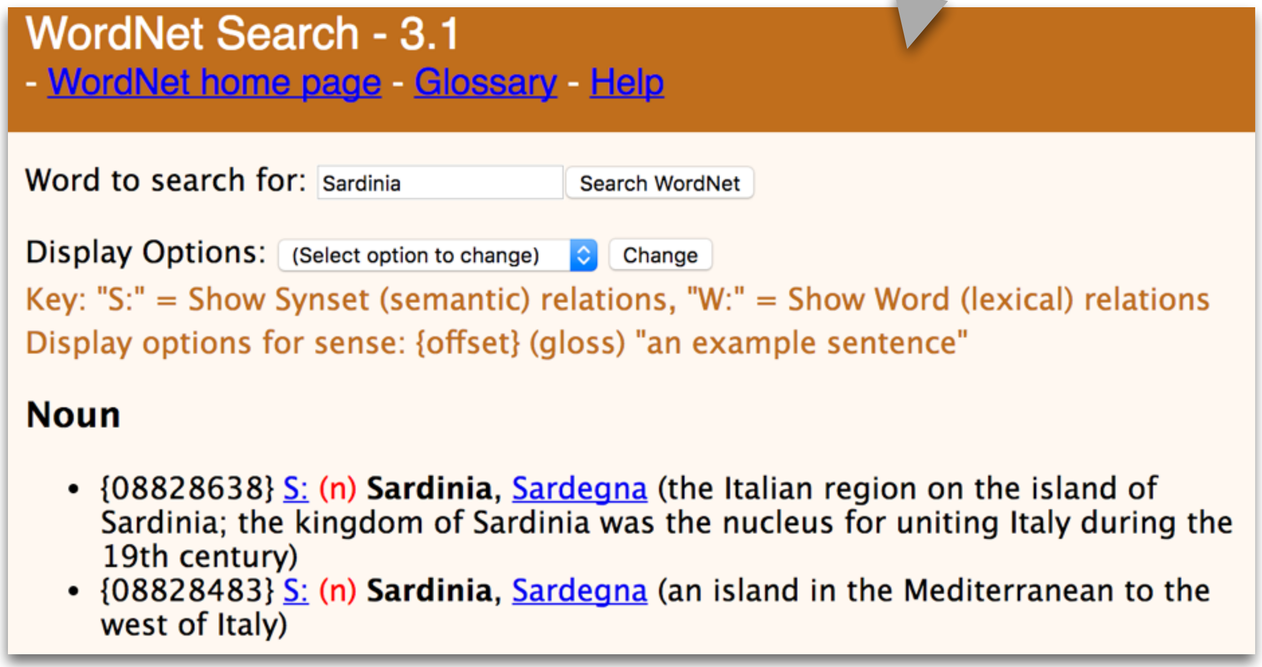
EL 18 de octubre de 2010 Xi fue nombrado vicepresidente de la Comisión Militar Central, de esta manera pasó a ocupar puestos destacados en el Ejército Popular de Liberación, el Partido Comunista de China y el Gobierno.

El 14 de marzo de 2013, tras la confirmación de la Asamblea Popular Nacional de China, Xi asumió la presidencia de la República Popular China.

Véase también [editar]

Dictionaries

- Link a symbol to its senses
- Sense: unique ID grouping all synonyms
  - for instance: **synsets** of WordNet, BabelNet



Linking and Disambiguation

- Two major problems with dictionaries
- Linking: identify the parts of text that must be linked to a certain concept/entity
  - e.g.: "*provincia di Cagliari*" vs. "*provincia*" di "*Cagliari*"
- Disambiguation: which sense to pick among a set of possible senses
  - e.g.: "calcio": sport, kick, calcium or part of a weapon
  - It is not possible to disambiguate without knowing the **context**

Senses and context

- “Sarà la **Vecchia Signora** a volare a Cardiff”

?

Context

- Context determines the sense
- (*Wittgenstein : we understand words if we learn to use them*)
- Distributional Hypothesis: similar words tend to appear in similar contexts
- Word2Vec [Mikolov et al., Distributed Representations of Words and Phrases and their Compositionality]
- Latend Semantic Analysis

Comparing sentences

- Word2Vec is a good tool to compare the sense of words
- Comparing whole sentences can be really difficult
  - Is context the same? Is it the same event?
    - "Prices jumped overnight" – "Two people jumped overnight in the North End"

Figurative use

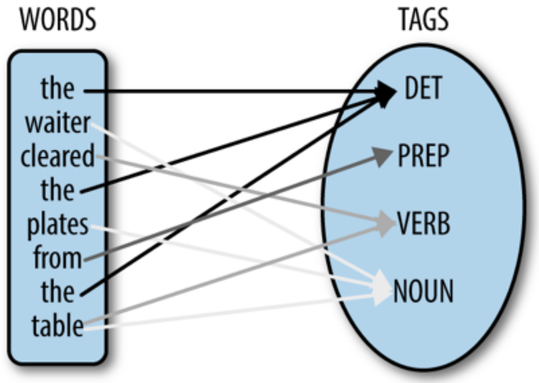
Ambiguity

  - Are emotions the same?
    - "I love when it's sunny:)" – "I love when it rains :("

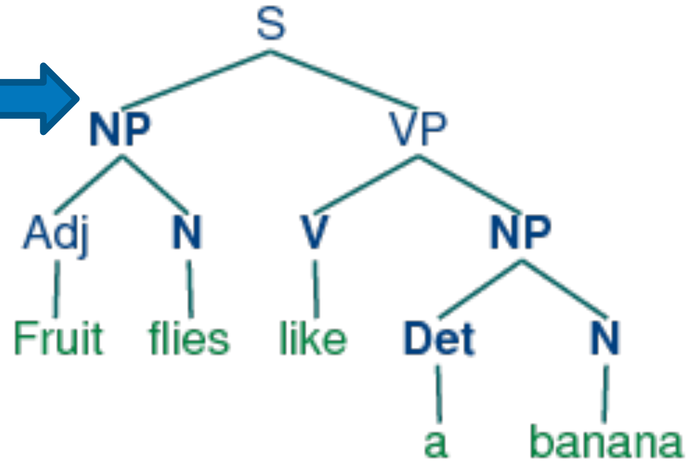
Irony

Evolution of Natural Language Processing

- POS Tagging



- Syntactic Analysis



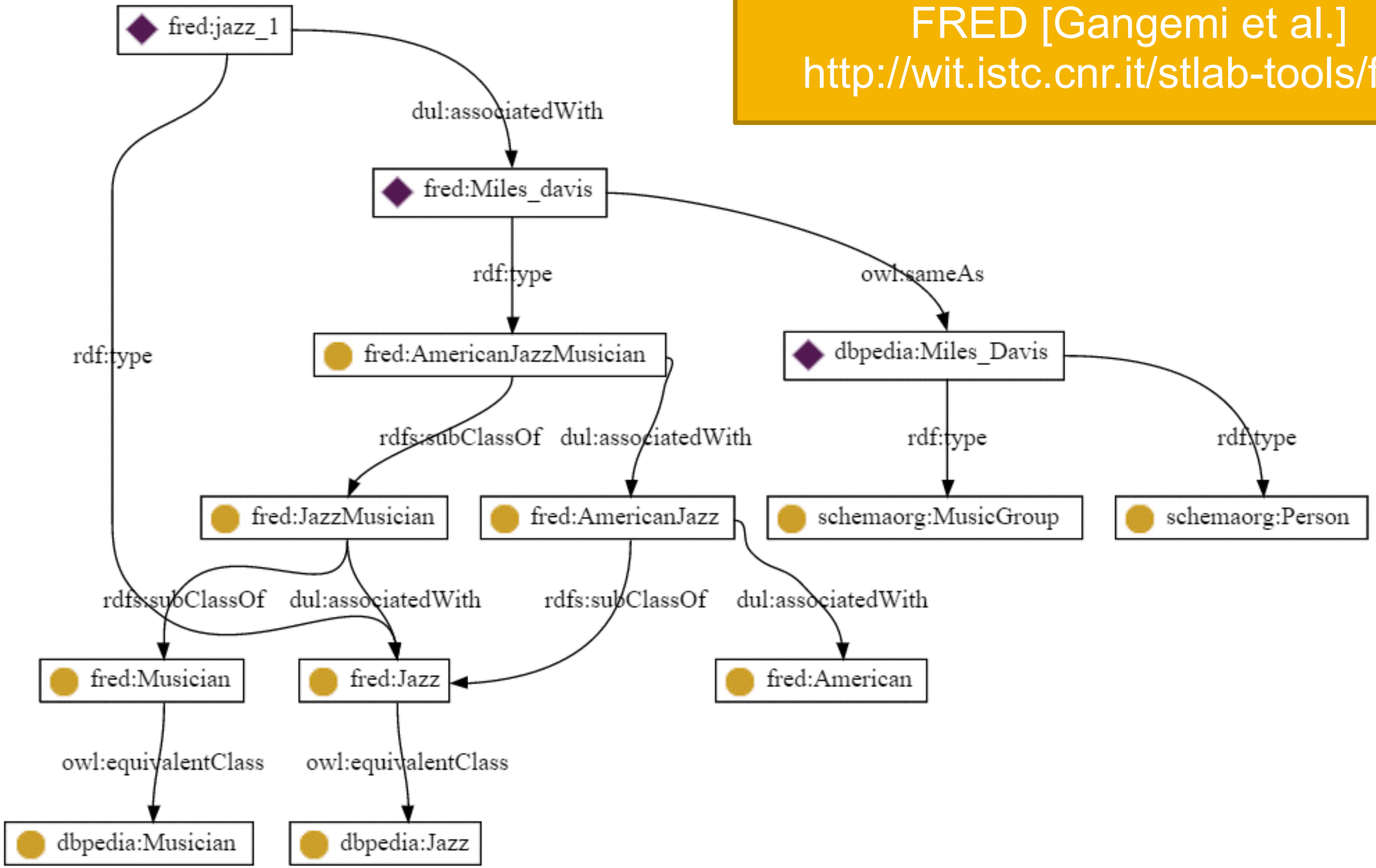
- Semantic Analysis

?

Babelfy [Navigli et al. 2013]

“Miles Davis was an American jazz musician”.

FRED [Gangemi et al.]
http://wit.istc.cnr.it/stlab-tools/fred/



Ontologies

- To convert a sentence in a semantic graph we need concepts and the relations between them
- **Ontologies** are lexical resources describing a domain with concepts, their properties and relationships
- They need a lot of work from experts
  - -> Great interest in methods to (semi-)automatically build or enrich ontologies

Beyond semantics

- Human language is complicate: we already saw irony and ambiguity;
- Information vs. opinions
- For instance:
  - “Trump is Putin’s dog”
  - “Donald Trump is the current president of the USA”

Subjective, negative opinion

Objective (fact)

From sense to emotions: Sentiment Analysis

Opinion Mining vs. Sentiment Analysis

- **Opinion mining** is more oriented to finding what people think *about a product, service or political actor*
- **Sentiment Analysis** is more oriented to finding *what people think*
  - In any case: interchangeably used, since methods are the same
- A definition by B.Liu :
  - “Computational study of opinions, sentiments, subjectivity, evaluations, attitudes, appraisal, affects, views, emotions, etc., expressed in text”

Opinions

- Subjective statement about something (a product, a situation, a service, a topic, etc…)
- They are key in e-commerce since they represent the word-of-mouth that allows potential buyers to convince themselves
  - So opinions are ubiquitous on the web

Products (amazon.com)

Customer Reviews

13,773

4.6 out of 5 stars

5 star		10,748
4 star		1,952
3 star		499
2 star		209
1 star		365

Share your thoughts with other customers

Write a customer review ›

See all 13,773 customer reviews ›

Top Customer Reviews

Well worth buying.

By Amazon Customer on 21 April 2017

Configuration: With Special Offers | Connectivity Technology: Wi-Fi Only | Colour: Black | Verified Purchase

Small and light enough for a handbag so perfect for long journeys. The screen is glare free and very easy to read no matter what the lighting conditions are. Easy to use, easy to charge and perfect for those who can't hold heavy books.

Comment | 14 people found this helpful. Was this review helpful to you? Yes No Report abuse

Give me convenience, or give me death

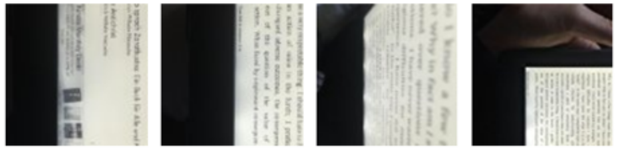
By I. Dransfield on 12 May 2017

Configuration: With Special Offers | Connectivity Technology: Wi-Fi Only | Colour: Black | Verified Purchase

'No more faff with having to carry books around' - that was my main reason for wanting a Kindle. I have a commute, like a few of us in the world do, so I want convenience. History, tradition, delicious smell of real books be damned. So I grudgingly took the plunge. And I'm very happy I did.

That main reason - not wanting to stuff my bag full of unwieldy paper monsters for my to-and-from work commute - immediately proved to be right on the money; the perfect reason to pick up a Kindle. I now have hundreds of books - more than I even have back at home - being carried around on this one, relatively slim, piece of kit that fits easily in my bag.

Customer Images



See all customer images

Most Recent Customer Reviews

Four Stars

It is lovely

Published 7 minutes ago by Mrs B Barrett

Five Stars

Easy to set up, neat and clear print and screen

Published 2 hours ago by Amazon Customer

70 year old says better than a book

Excellent.Much better than paper books.

Published 3 hours ago by Amazon Customer

I upgraded to their Kindle and like it better than the original one

I upgraded to their Kindle and like it better than the original one. It is lighter and more compact, the screen is easier to read and the backlight function is perfect. Read more

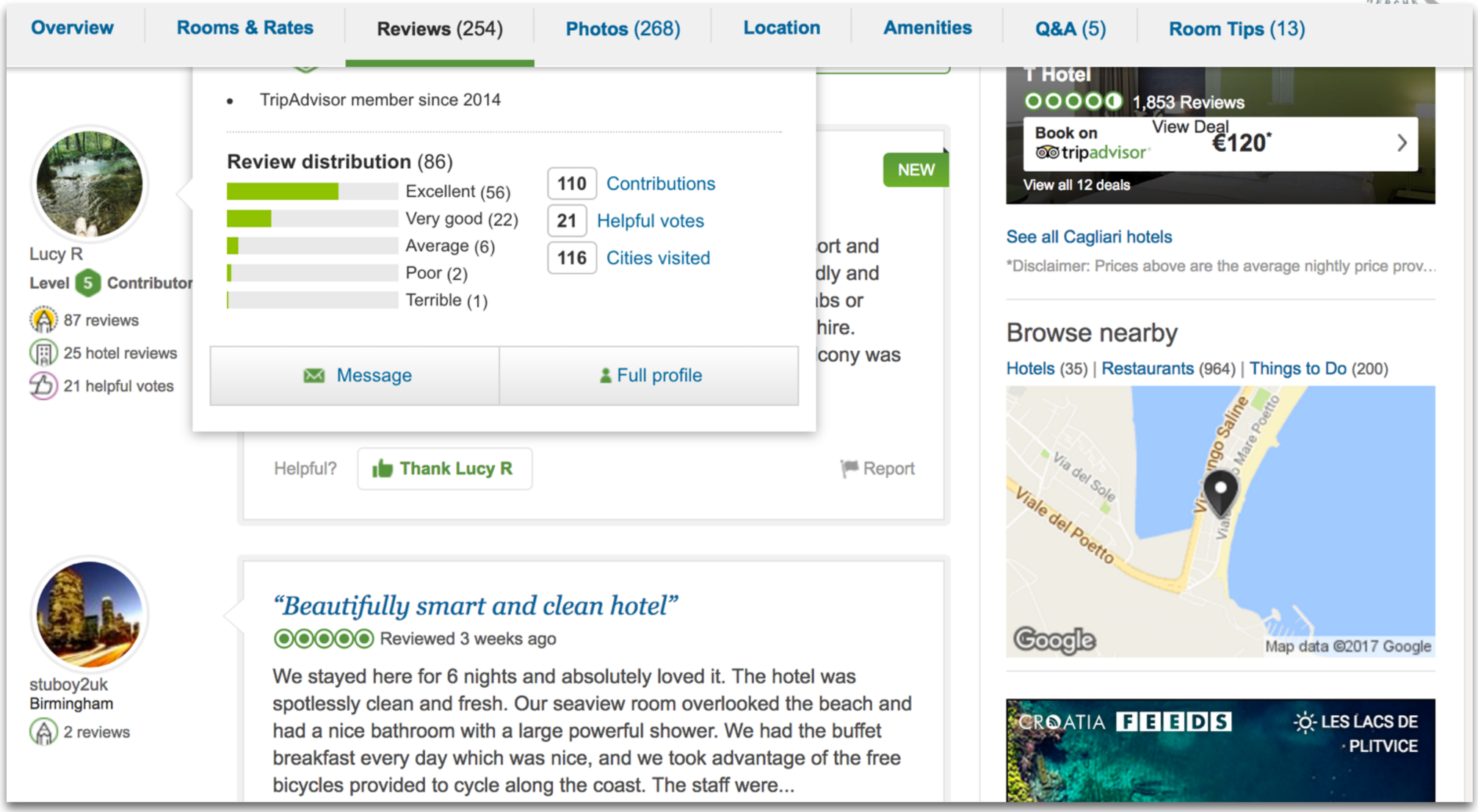
Published 3 hours ago by john

As good as the original, just a bit different

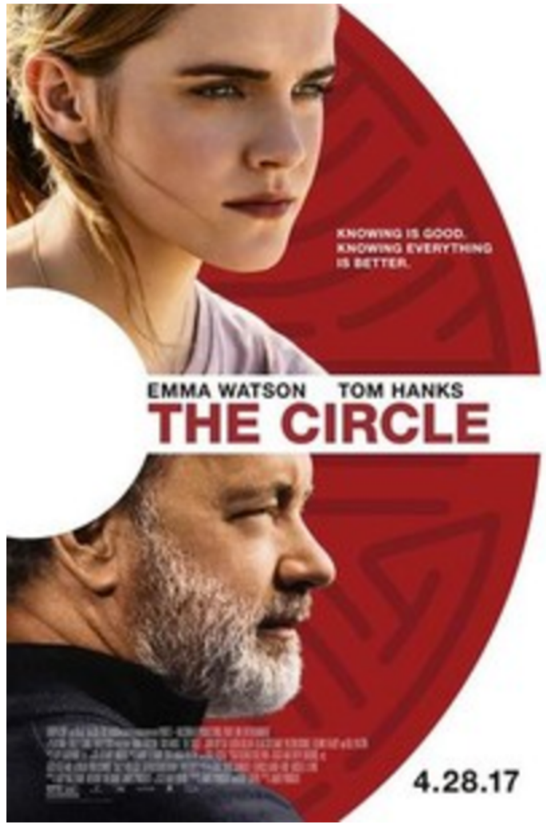
This is my second Kindle, the first being one of the earlier ones with keyboard & buttons to move the pages on. Read more

Published 4 hours ago by Milo

Travels (TripAdvisor)



Movies (Rotten Tomatoes)



TOMATOMETER

All Critics | Top Critics

16%

Average Rating: 4.3/10
Reviews Counted: 86
Fresh: 14
Rotten: 72

Critics Consensus: *The Circle* assembles an impressive cast, but this digitally driven thriller spins aimlessly in its half-hearted exploration of timely themes.

AUDIENCE SCORE

24%
liked it

Average Rating: 2.3/5
User Ratings: 13,616

ADD YOUR RATING

+ WANT TO SEE

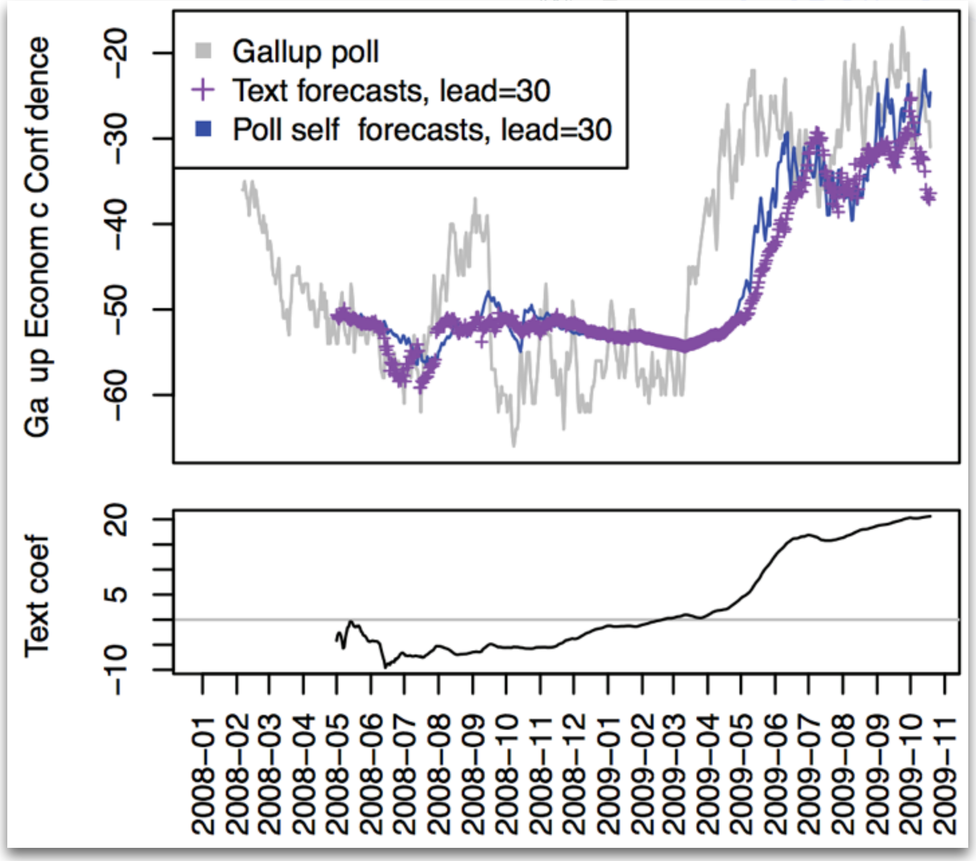
NOT INTERESTED

Add a Review (Optional)

Post

Politics, etc. (Twitter)

What's the score here?



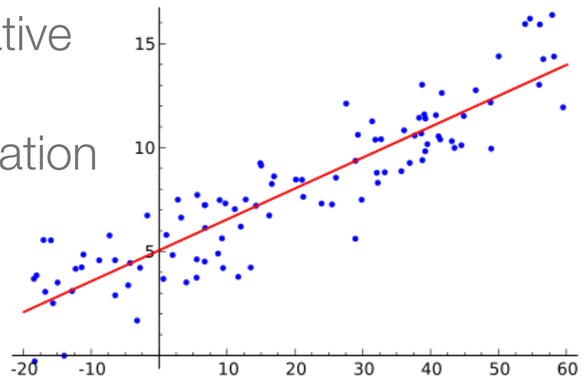
O'Connor et al., 2010:
"From tweets to polls: Linking text sentiment to public opinion time series"

Sentiment Analysis in text

- D.Jurafsky: Sentiment Analysis is the detection of *attitudes*:
    - Attitudes: “enduring, affectively colored beliefs, dispositions towards objects or persons”
- Holder (source) of attitude
- Target (aspect) of attitude [in opinion mining]
- Type of attitude
    - From a set of types : Like, love, hate, value, desire, etc.
    - Or (more commonly) simple weighted polarity: positive, negative, neutral, together with strength
- Granularity:
    - Entity or feature/aspect, sentence, or whole document level

Sentiment Analysis in text

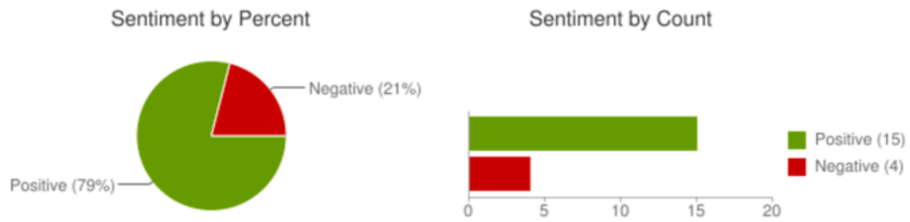
- (B.Liu) Usually consists of two steps:
  - Subjectivity classification (To identify subjective sentences)
  - Sentiment classification of subjective sentences into two classes, positive and negative
    - Also: prediction of ratings ★-> ★★★★★ : use **regression** instead of classification



- Caveats:
  - Many objective sentences can imply sentiments
    - E.g., “Russia, a victim of the cyberattack, voices outrage”
  - Many subjective sentences do not express positive or negative sentiments/opinions
    - E.g.,”I believe he went home yesterday.”

Challenges

- Examples: output from http://www.sentiment.140



Tweets about: Macron

caymsortweeted: RT @JeremyCliffe: Eurozone recovering, **Macron** elected, refugee crisis off the headlines... perhaps Cameron should have held the Brexit vote?
Posted: 17 seconds ago

?

LaCrampe: RT @jbarro: Emmanuel **Macron** is 39 and is the president of a G7 country. What have you done with your life?
Posted: 21 seconds ago

JLDERCLE: RT @V_of_Europe: How sweet.. new French president **Macron** kissed his grandma today! https://t.co/BCN5WBBEtS
Posted: 24 seconds ago

?

DEPLORABLERAY11: RT @OliMauritania: Like a typical Democrat in the US, **Macron** talks a lot about "diversity" on campaign trail but a picture of his staff say?
Posted: 27 seconds ago

?

NabilAbdullah20: RT @BBCBreaking: Emmanuel **Macron** inaugurated as French president, the country's youngest leader since Napoleon https://t.co/cHTIFlRsOA htt?
Posted: 41 seconds ago

CTJaddie: RT @altrockchick1: The US is still too partisan; neither side trusts the other. You need a **Macron** to emerge to move this beyond partisan po?

Challenges

Galaxy Note 7

JulianWaynsen: **Galaxy Note 7** Recall: What you need to know https://t.co/3kVS4iXkBX via @YouTube
Posted: 30 minutes ago

roshd: RT @dave_cactus: You can have my Samsung **Galaxy Note 7** when you pry it from my hot scorched hand.
Posted: 31 minutes ago

Minorthree: RT @chinhdo: Great. The fingerprint scanner on my brand new @Samsung **Galaxy Note 7** decided to quit - nope, restarts didn't fix it https://t?
Posted: 43 minutes ago

X

BateroElias: **Galaxy Note 7** Drop Test from 1,000 FEET!! DID IT SURVIVE?? https://t.co/UYTkzKqN5x vía @YouTube
Posted: 51 minutes ago

?

GadgetRumours: G+: U.S. urges Samsung **Galaxy Note 7** owners to stop use, return after 26 burned, 55 property damage fires? https://t.co/nr9ydGcizh
Posted: 52 minutes ago

amzeratul: @Olifiers @quickfingerz @matkins85 You just need an iPhone **7** to film underwater, and a **Galaxy Note 7** to blow a hole in the ice. :D
Posted: 56 minutes ago

X?

TruPlaya1870: @TELUSsupport ok thanks for that info! Could you tell me which version should I have in my **galaxy Note 7**?
Posted: 1 hour ago

Hoopgroundzero: RT @androids: **Galaxy Note 7** battery replacement receives new complaints - https://t.co/22xHDtjk2M https://t.co/MF06bfJiSj
Posted: 1 hour ago

OfficialMiLKK: @SamsungMobileUS When will more **Galaxy Note 7** devices be available because there are none left where I live in Texas currently :(
Posted: 1 hour ago

Challenges

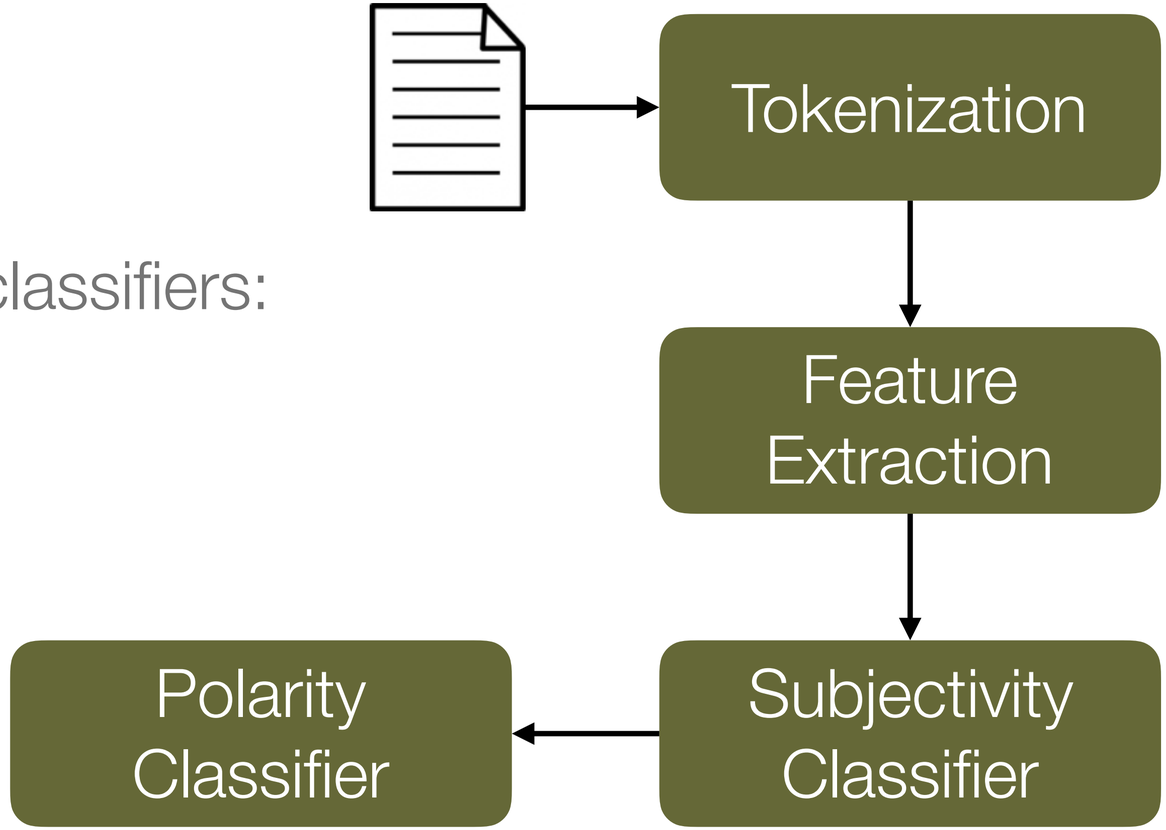
- Subjectivity detection: tell an opinion from a fact
- Handling negations
  - “I expected to be disappointed with the device. This was not the case.”
- Context switches
  - “I loved this movie when I was little. Now I hate it!”
- Irony
  - “I love when I'm ready to go to sleep but can't because I can clearly hear my neighbors music from inside my house!”

Data

- SNAP Amazon reviews with ratings:
  - https://snap.stanford.edu/data/web-Amazon.html
- Cornell U. Movie reviews:
  - http://www.cs.cornell.edu/people/pabo/movie-review-data/
- SemEval
  - http://alt.qcri.org/semeval2015/task10/ comprehensive task
  - http://alt.qcri.org/semeval2015/task11/ dedicated to figurative language
- Italian
  - SENTIPOLC16 http://www.di.unito.it/~tutreeb/sentipolc-evalita16/
  - SENTIPOLC14 http://www.di.unito.it/~tutreeb/sentipolc-evalita14/
- French
  - DEFT2015 https://deft.limsi.fr/2015/ , DEFT2017 https://deft.limsi.fr/2017/
- Spanish
  - TASS2015 http://www.sepln.org/workshops/tass/2015/tass2015.php

A baseline method

- Tokenization
- Feature Extraction
- Classification using different classifiers:
  - Naïve Bayes
  - MaxEnt
  - SVM
  - …

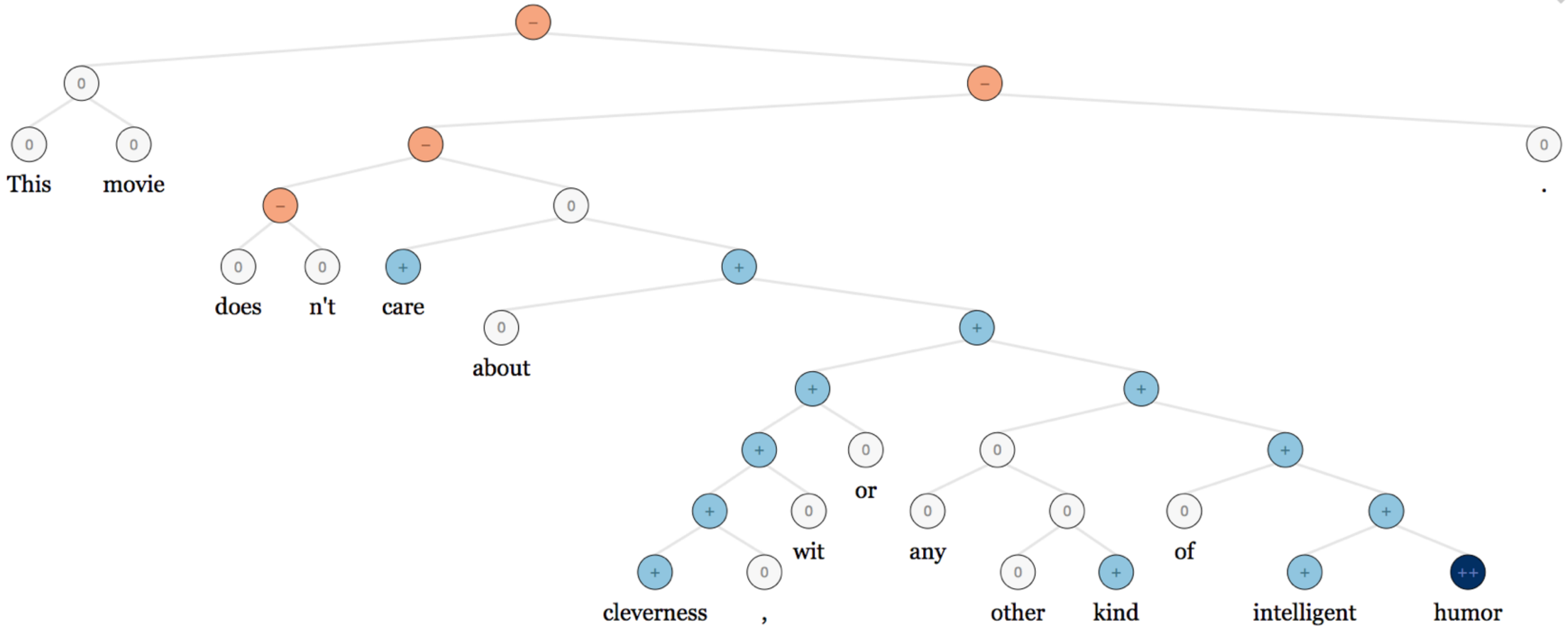


Tokenization

- Splitting text into **tokens** - various decisions:
  - word tokens? “dare”, “una”, “mano”
  - n-grams?
    - word n-grams : “dare una mano” (3-grams)
    - character n-grams: “dar” “are” “#Re” “Ren” “enz” (3-grams)
  - Punctuation
  - Dependency parsing?

A dare una mano a #Roma, Mr. #Renzi, dovevate cominciare quando @ignaziomarino era sindaco. http://bit.ly/o2er5ty
…
Sta superando "Grillo" !Lo vedremo presto a Zelig 😂

Dependency parsing



http://nlp.stanford.edu:8080/sentiment/rntnDemo.html

Extracting features

- Selection of tokens to be used as features in the classifier
  - Lemmatization: “dovevate” -> “dovere”
    - Stemming: “dov-” ?
    - Removing stopwords?
  - Normalization to lowercase?
  - Normalization of web addresses?
  - Linking entities?
    - (e.g. #Renzi is the same as @matteorenzi and the same as http://dbpedia.org/page/Matteo_Renzi)
  - Emojis, smileys
  - Negations: “didn’t like the movie” -> “not-like” “not-the” “not-movie”
  - POS-tagging (to select adjectives, nouns, pairs adjective-noun, etc.)
  - Filter by frequency?

A dare una mano a #Roma, Mr. #Renzi, dovevate cominciare quando @ignaziomarino era sindaco. http://bit.ly/o2er5ty
…
Sta superando "Grillo" !Lo vedremo presto a Zelig 😂

Examples of features

- Evalita-14

Subjectivity task

Word	SUBJ	p(w|SUBJ)	NOT SUBJ	p(w|NOT SUBJ)	p(W|SUBJ)-p(W|NSU
!	1048	0,024144128	104	0,005336617	0,01880751
non	602	0,01386905	123	0,006311576	0,007557474
governo	692	0,015942496	212	0,010878489	0,005064007
...	593	0,013661706	180	0,009236453	0,004425253
mi	198	0,004561581	14	0,000718391	0,003843191
ma	263	0,00605907	46	0,002360427	0,003698643
)	279	0,006427683	63	0,003232759	0,003194924
si	249	0,005736534	69	0,00354064	0,002195894
se	163	0,003755241			
più	171	0,003939548			
tutto	116	0,002672442			
ci	148	0,003409667			

Polarity task

Word	POS	p(w|POS)	NEG	p(w|NEG)	diff prob
grazie	59	0,003125828	16	0,000527861	0,002597967
)	158	0,008370861	128	0,004222889	0,004147972
D	24	0,001271523	7	0,000230939	0,001040584
amo	20	0,001059603	3	0,000098974	0,000960629
<3	19	0,001006623	2	6,60E-05	0,00094064
Grazie	18	0,000953642	4	0,000131965	0,000821677
piace	32	0,001695364	19	0,000626835	0,001068529
te	18	0,000953642	6	0,000197948	0,000755694
Buon	12	0,000635762	2	6,60E-05	0,000569779

Classifying using Naïve Bayes

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(w_i \mid c_j)$$

$$\hat{P}(w \mid c) = \frac{count(w, c) + 1}{count(c) + |V|}$$

- Features are usually *binarized*: experiments suggest that having a feature more than one time in a text doesn't add any information to just having it once
- SVM, MaxEnt give better results (SVM better for high dimensional spaces)

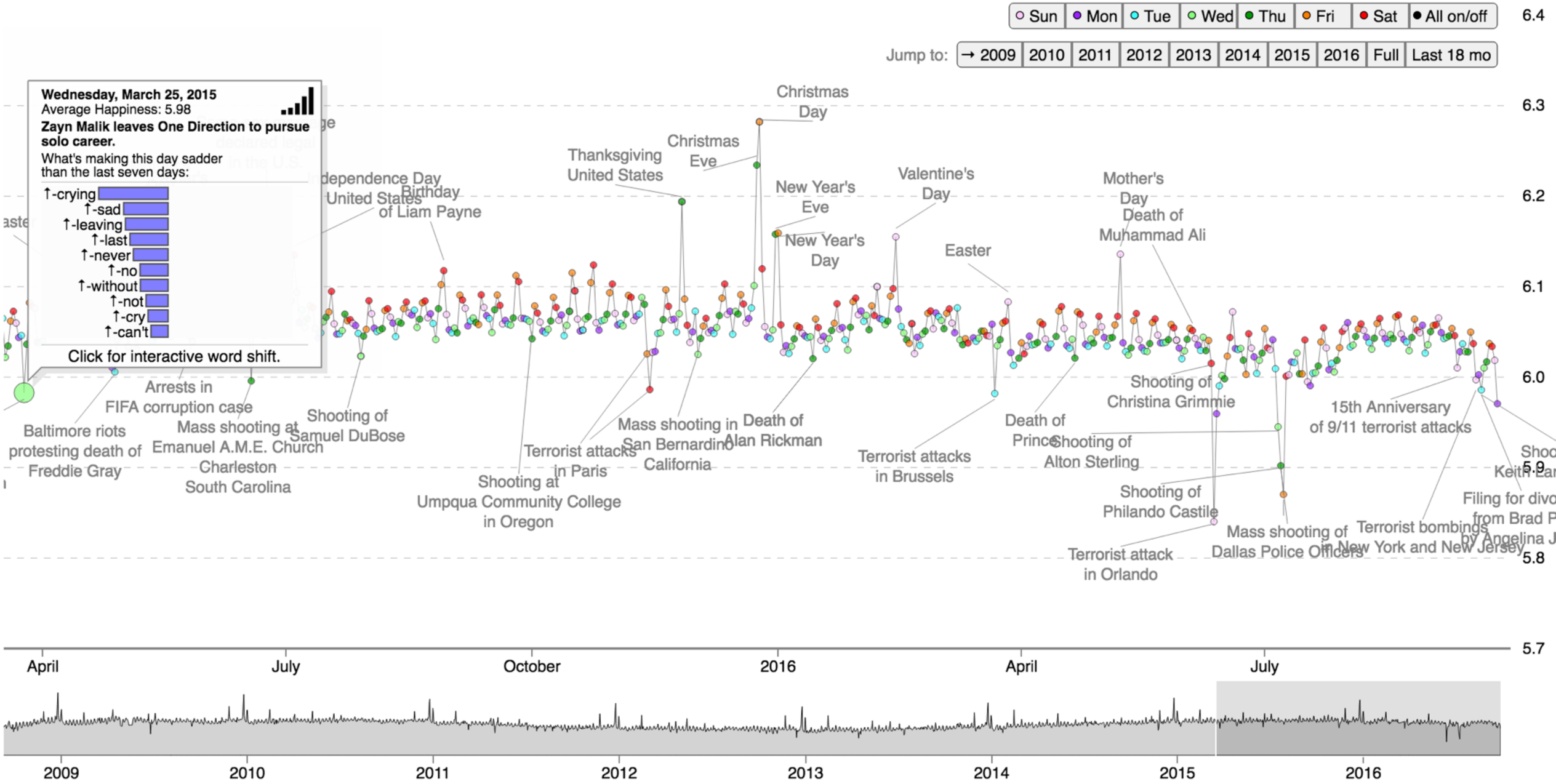
Sentiment Lexicons

Sentiment Knowledge

- While dictionaries map words into senses, sentiment lexicons map words (or expressions) into sentiments:
  - word -> polarity (positive, negative)
  - word -> emotion (joy, sadness, fear…)
- Sentiment lexicons are very effective in predicting polarities
  - Example: hedonometer.org, labMT lexicon

Average Happiness for Twitter

hedonometer.org



labMT

- Downloadable from: https://github.com/andyreagan/labMT-simple/tree/master/labMTsimple/data/LabMT
- 10,223 words with assigned scores
  - Scores (float) from 1 to 9
- Multilingual (Machine Translated)

General Inquirer

- Home page: http://www.wjh.harvard.edu/~inquirer
- List of Categories: http://www.wjh.harvard.edu/~inquirer/homecat.htm
  - Categories:
    - Positiv (1915 words) and Negativ (2291 words)
      - Strong vs Weak, Active vs Passive, Overstated versus Understated
      - Pleasure, Pain, Virtue, Vice, Motivation, Cognitive Orientation, etc
- Free for Research Use

LIWC (Linguistic Inquiry and Word Count)

- Home page: http://www.liwc.net/
- 2300 words, >70 classes
- Affective Processes
  - negative emotion (bad, weird, hate, problem, tough)
  - positive emotion (love, nice, sweet)
- Cognitive Processes
  - Tentative (maybe, perhaps, guess), Inhibition (block, constraint)
- Pronouns, Negation (no, never), Quantifiers (few, many)
- Fee

MPQA

- Home page: http://www.cs.pitt.edu/mpqa
- 6885 words from 8221 lemmas
  - 2718 positive
  - 4912 negative
- Each word annotated for intensity (strong, weak)
- GNU GPL

AFINN

- Home page: https://github.com/wooorm/afinn-165
- 3382 entries
  - Range: -5, +5 (discrete)

B.Liu opinion lexicon

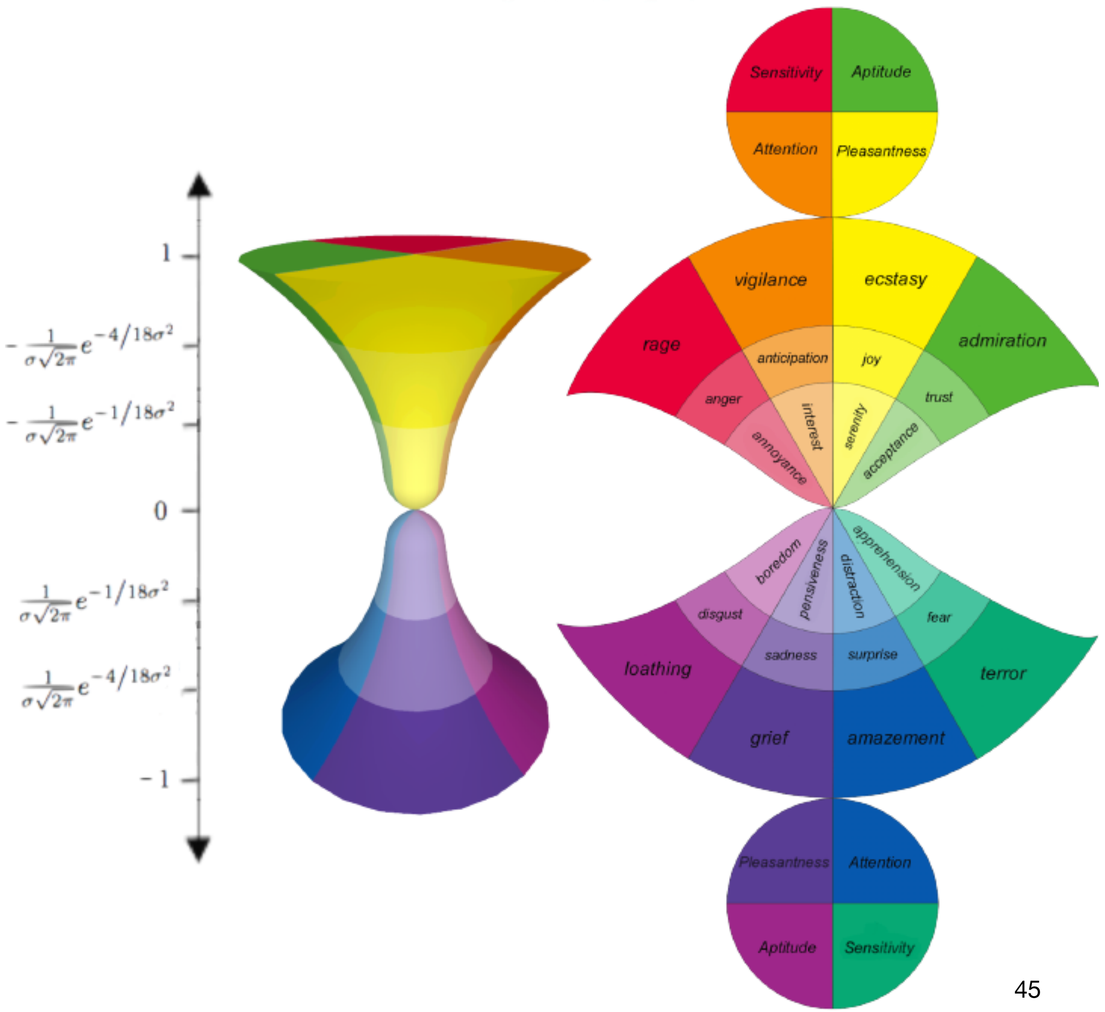
- Bing Liu's Page on Opinion Mining
- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar
  - 6786 words
  - 2006 positive
  - 4783 negative
- Binary polarity -1, +1, lemmatized words

SentiWordNet

- Home page: http://sentiwordnet.isti.cnr.it/
- All WordNet *synsets* automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
  - [estimable(J,3)] “may be computed or estimated”
    - Pos 0 Neg 0 Obj 1
  - [estimable(J,1)] “deserving of respect or high regard”
    - Pos .75 Neg 0 Obj .25

sentic.net

- http://sentic.net/
- 50,000 concepts with an intensity assigned
- -1 is extreme negativity and +1 is extreme positivity
- Different “axes” of emotions:
  - rage <-> terror
  - admiration <-> loathing
  - ecstasy <-> grief
  - vigilance <-> amazement



NRC Affect Intensity Lexicon

- Home page: http://saifmohammad.com/WebPages/AffectIntensity.htm
- Similarly to sentic.net , different emotions:
  - anger, fear, joy, sadness
  - real values from 0 (not associated) to 1 (strongly associated)
- Around 6000 entries

Using Sentiment Lexicons

- Problems:
  - **Weighting** (especially for conflicting scores)
    - In SentiWordNet, you can have different scores for different senses of a word
    - Specific case: translation into another language
  - Negation
    - Solution (Taboada et al.): inverting score, or shifting score towards the other end
  - Meaning of conflicting scores in the same sentence
    - Ex: “My phone is shattered as well my hopes and dreams”

-2 +2 +1

Multilingual lexicons

- Two strategies:
  - Bootstrapping from known polarized words and expand with similarity (WordNet, semantic vectors)
    - Example: **ItEm** (Passaro et al., UniPi)
  - Automatically translate and propagate the score to all of its translations
    - disturbing (-2) -> inquietante (-2), sconvolgente (-2), turbato (-2)
    - If a score already existed for the translation, keep the most polarizing value
    - Repeat for all words
    - Example: **FEEL** (French): http://www.lirmm.fr/~abdaoui/FEEL.html

Some insights about IRADABE

In collaboration with Irazú Hernandez-Farias (UPV, Università di Torino)

IRADABE: a lexicon-enhanced supervised system for Italian sentiment analysis

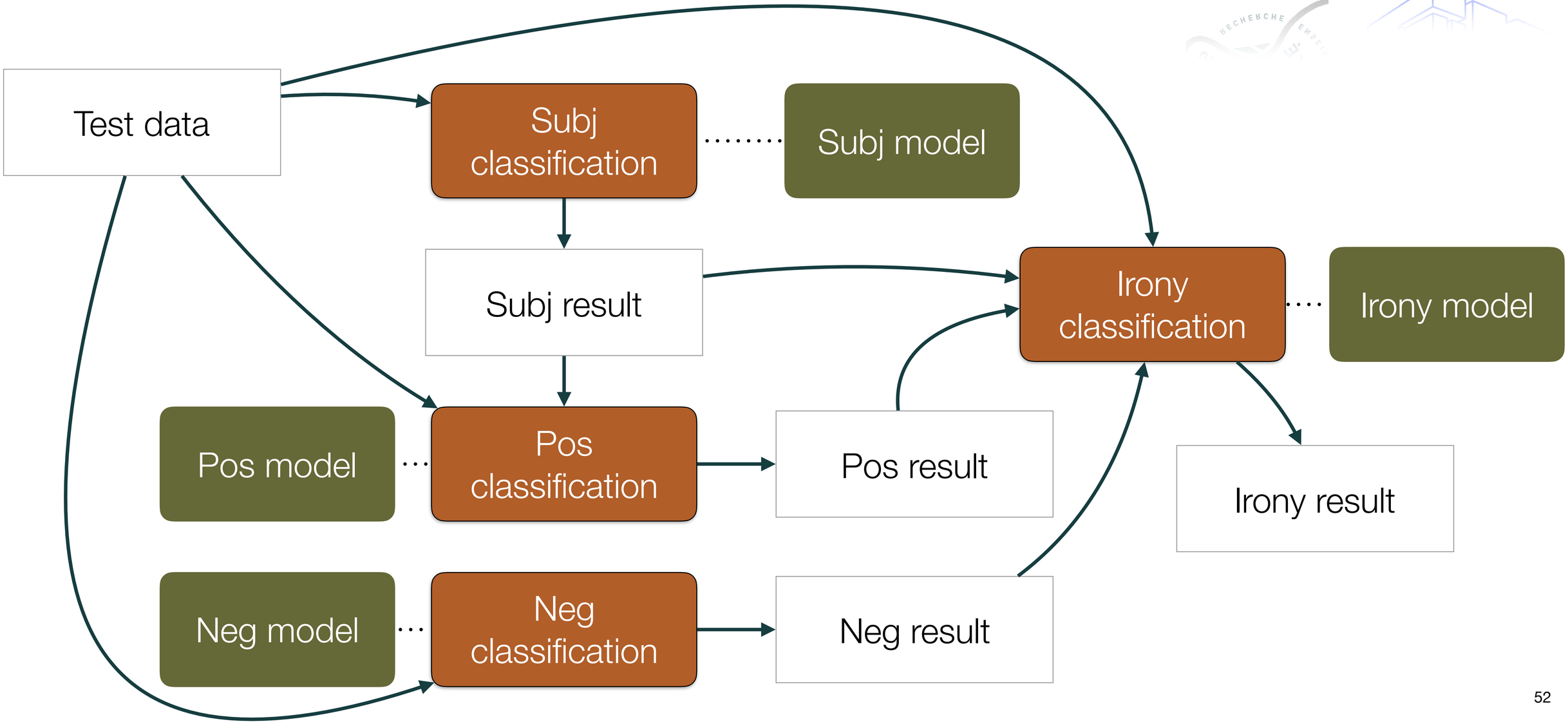
- Features:
  - positional BOW, smileys, uppercase/lowercase ratio and lexicon-derived features…
  - Detection of http addresses, personal pronouns, adversative connectors and negations
  - For polarity and irony: result of subjectivity classification (if a comment is objective, it cannot be positive nor negative)
- Classification method:
  - Support Vector Machine (LibSVM)
    - https://www.csie.ntu.edu.tw/~cjlin/libsvm/
  - RBF kernel

The Italian SENTIPOLC Task

subj	pos	neg	iro	description
0	0	0	0	an objective tweet
1	0	0	0	a subjective tweet with neutral polarity and no irony
1	1	0	0	a subjective tweet with positive polarity and no irony
1	0	1	0	a subjective tweet with negative polarity and no irony
1	1	1	0	a subjective tweet with both positive and negative polarity (mixed polarity) and no irony
1	1	0	1	a subjective tweet with positive polarity, and an ironic twist
1	0	1	1	a subjective tweet with negative polarity, and an ironic twist

0	0	0	0	l'articolo di Roberto Ciccarelli dal manifesto di oggi `http://fb.me/1BQVy5WAk`
1	0	0	0	Primo passaggio alla #strabrollo ma secondo me non era un iscritto
1	1	0	0	splendida foto di Fabrizio, pluri cliccata nei siti internazionali di Photo Natura `http://t.co/GWoZqbxAuS`
1	0	1	0	Monti, ripensaci: l'inutile Torino-Lione inguaia l'Italia: Tav, appello a Mario Monti da Mercalli, Cicconi, Pont... `http://t.co/3CazKS7Y`
1	1	1	0	Dati negativi da Confindustria che spera nel nuovo governo Monti. Castiglione: "Avanti con le riforme" `http://t.co/kIKnbFY7`
1	1	0	1	Letta: sicuramente non farò parte del governo Monti . e siamo un passo avanti. #finecorsa
1	0	1	1	Botta di ottimismo a #lInfedele: Governo Monti, o la va o la spacca.

Architecture



Positional BOW

- Select from training set all the words with frequency > 5 and abs(*I*(*w*)) > 0.001:

$$I(w) = p(w|c^+)\left(\log(p(w|c^+)) - \log(p(w|c^-))\right)$$

- These words will constitute the BOW dictionary *D*

- The weight of a word *w* in a sentence *s* is calculated as:

$$W(w) = 1 + pos(w)/len(s)$$

  - Where pos(w) is the last position of the word in the sentence

- The reason is that we observed that when some “sentiment” words are used towards the end, their effect is increased

Sentiment Lexicons

- AFINN, Hu-Liu, SentiWordNet, labMT
- Need to translate the lexicons from English to Italian
  - Weight is propagated to all of its translations
    - *disturbing* (-2) -> *inquietante* (-2), *sconvolgente* (-2), *turbato* (-2)
  - In case of conflicts: keep the most polarizing score
- SentiWordNet: use the average of scores over all the senses of a word

Merging lexicons: UnISeLex

- Better **coverage**: some words are listed in some dictionaries but not in other ones - or they are not listed at all
  - Exemple: *sgradevole* in AFINN and SWN but not in Hu-Liu
  - *avarizia* is not present in any dictionary
  - Sizes (excluding multiwords):
    - AFINN 2242 words
    - Hu-Liu 3956 words
    - SWN 9531 words
- Solution: merging lexicons
- Final size for UnISeLex: 31601 words

Merging algorithm

- Create the starting dictionary D_0 by the union of all words in the dictionaries (12102)
  - The polarity score for each word is the average of non-zero values from the different dictionaries
- D_1 is created by extending the scores to all synonyms (using WordNet) (15412)
  - If the synonym has already a score, keep the most polarizing one (if same sign), or sum them (if different sign)
- D_f is created by extending the scores to the pseudo-synonyms (using Word2Vec) (31601)
  - vectors calculated on ItWaC (http://wacky.sslmit.unibo.it/)
  - If the pseudo-synonym has similarity > 0.6 and has not already a score, set it to the score of the originating word multiplied by the similarity value

Lexicon-derived features

- Options to transform a sequence of scores in an aggregate one:
- Sum of positive scores
- Sum of negative scores
- Count of positive scores
- Count of negative scores
- Polarity Balance (positive-negative)
- Average of scores
- Standard deviation of scores
  - to detect sudden changes of polarity – may be a clue to irony and/ or subjectivity:
    - « Saranno bravi eh, ma brutti...»

“My phone is shattered as well my hopes and dreams”

-2 +2 +1

3

-2

2

1

1

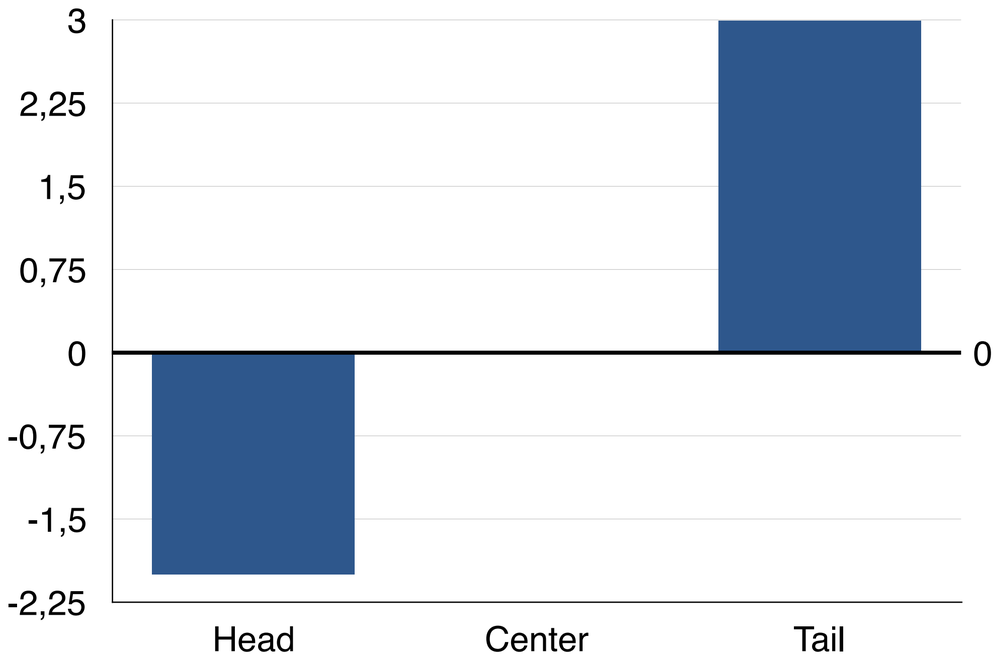
0.33

2.08

Relative Positional Lexicon weight

"My phone is shattered | as well my | hopes and dreams"

-2 +2 +1



IRADABE: best features

training set

	subj	pos	neg	iro
1	*http*	*subj*	*subj*	*subj*
2	*shout*	*AvgLabMT*	*Sum-*	*http*
3	*Sum-*	“grazie"	*Count-*	“governo"
4	*Count-*	*emo*	*AvgLabMT*	*mixed_pol*
5	*HeadS*	*PolBalance*	*sen_length*	*shout*
6	*pers*	*http*	*PolBalance*	“Mario"
7	“!”	“?"	*http*	“che”
8	*AvgLabMT*	*Sum+*	*MiddleS*	“#Grillo"
9	“mi”	“bello"	*taboo_words*	*sen_length*
10	*taboo_words*	“amare”	*StdDevLabMT*	*Sum-*

Legenda:

http: URL in tweet (positional score)
pers: personal pronouns
emo: smileys
shout: uppercase ratio
subj: subjective
Sum+/-: sum +/- scores
Count+/-: count +/- scores
PolBalance: polarity balance
HeadS: scores in the first part
MiddleS: scores in the central part
AvgLabMT: average of LabMT scores
StdDevLabMT: std. dev of LabMT scores

Results

- SentiPol 2014 (not the same feature set):

rank	team	F(C)
1	uniba2930	0.7140
2	UNITOR	0.6871
3	IRADABE	0.6706
4	UPFtaln	0.6497
5		0.5972
6		0.5901
7		0.5825
8	fbkshelldkm	0.5593
9	itagetaruns	0.5224
10	baseline	0.4005

Subj

rank	team	F(C)
1	uniba2930	0.6771
2	IRADABE	0.6347
3	CoLingLab	0.6312
4	UNITOR	0.6299
5		0.6049
6		0.6026
7		0.5980
8	fbkshelldkm	0.5626
9	mind	0.5342
10	itagetaruns	0.5181
11	Itanlp-wafi*	0.5086
12	baseline	0.3718
	*amended run	0.6637

Polarity

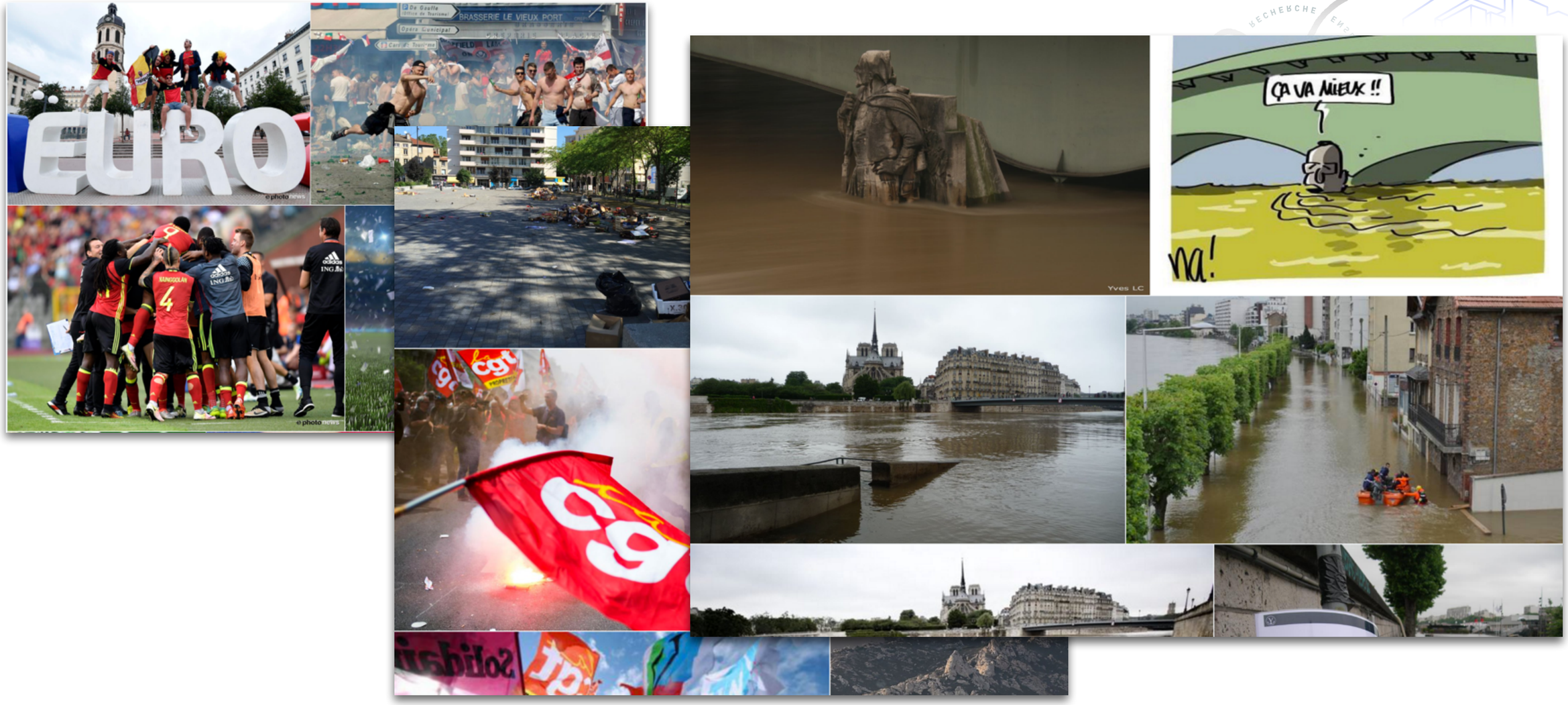
rank	team	F(C)
1	UNITOR	0.5759
2	IRADABE	0.5415
3	SVMSLU	0.5394
4	itagetaruns	0.4929
5		0.4771
6		0.4707
7		0.4687
8	baseline	0.4441

Irony

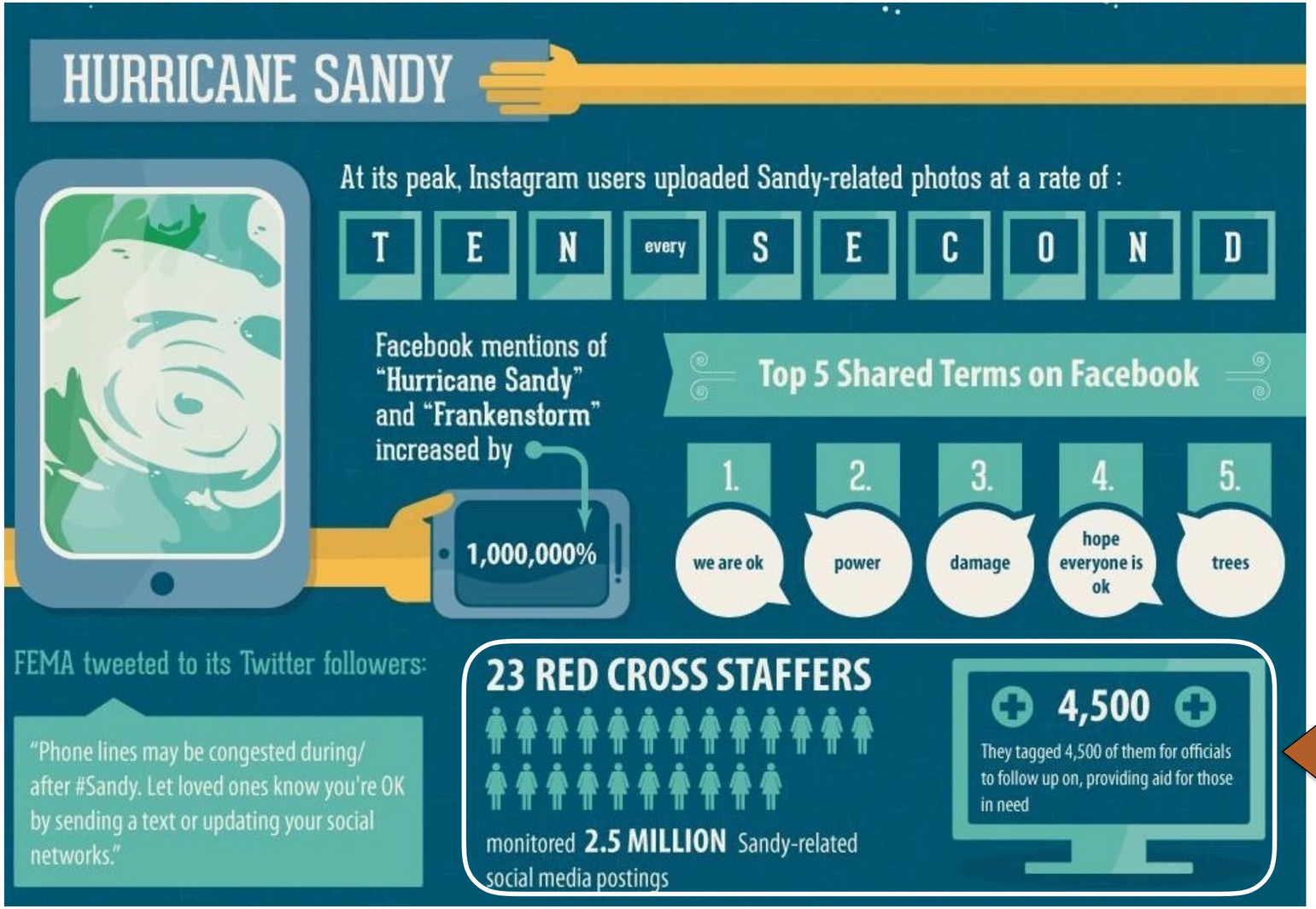
- SentiPol 2016:
  - subj:0.7108 (best 0.7444); polarity: 0.6453 (best 0.6643); irony: 0.5133 (best 0.5480)

Sentiment Analysis for Disaster Management

Tweet-boosting events



Social Networks and Disaster Management

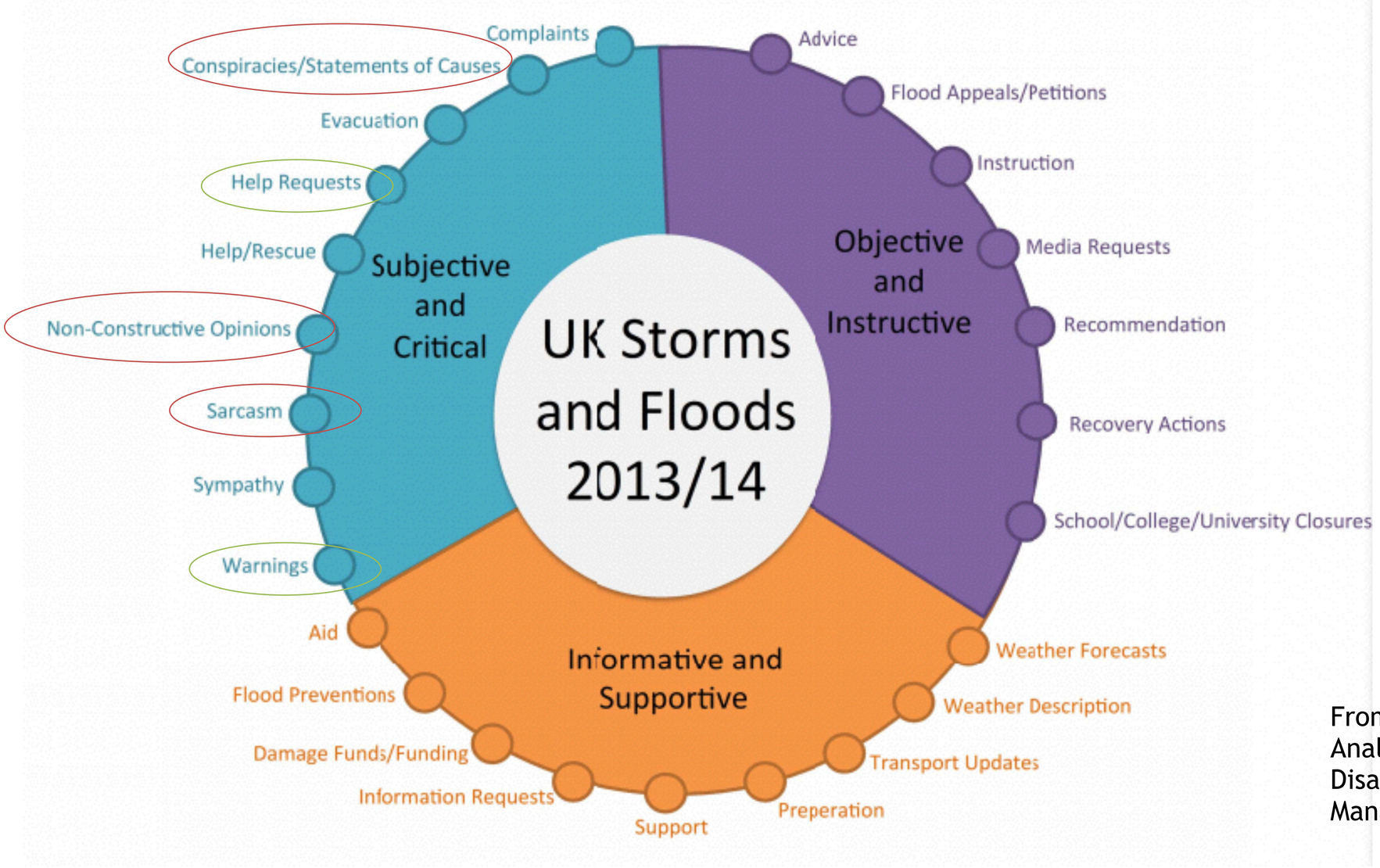


Top reasons to 'go social':

- Congested emergency lines
- One post to reach them all
- Unsure about the importance

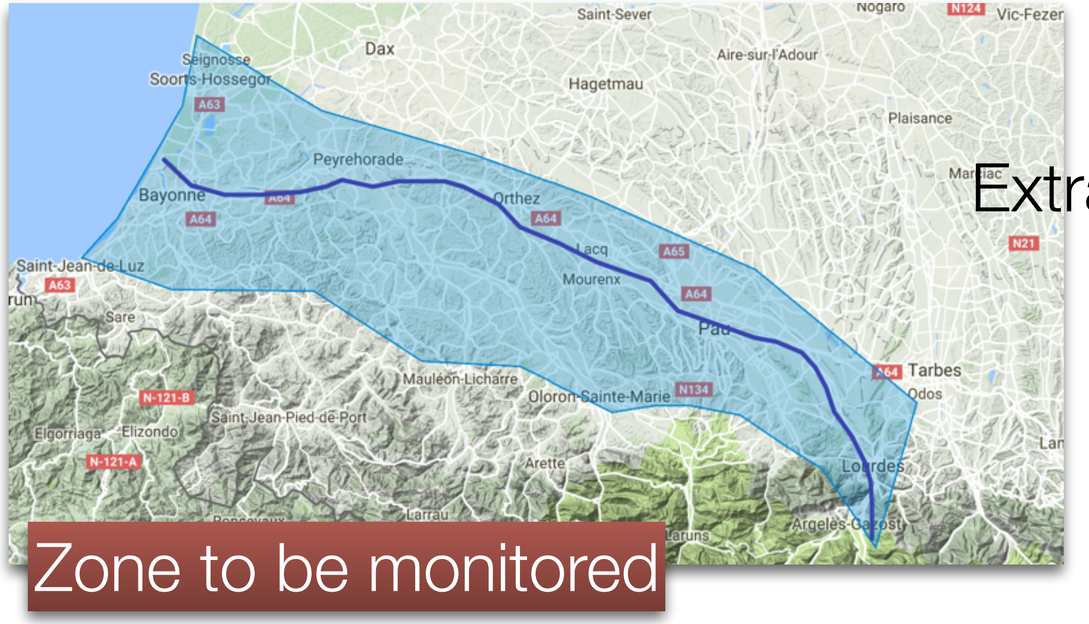
http://mashable.com/2013/05/21/social-media-disaster-response/#IKxYUNzazgq0

Social networks and Disaster Management



From: Parsons et al., Thematically Analysing Social Network Content During Disasters Through the Lens of the Disaster Management Lifecycle

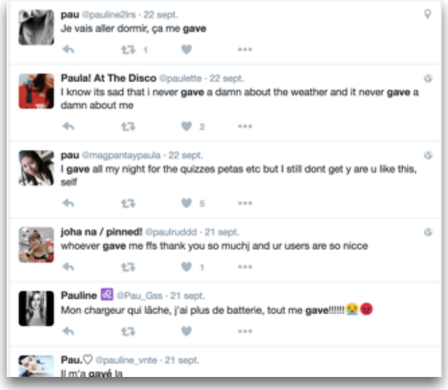
Project NaDiA-SMM



Zone to be monitored

Meteorological alert

Extracting related flux



Signalling relevant situations

Open issues

- Determine the location of the tweets
  - Only a small fraction of tweets are georeferenced wit lat, lon
  - Need to work on toponyms -> **ambiguity**
- Fine-grained Sentiment Analysis
  - Need to detect emotions (like fear)
- Reliability of Sentiment Analysis:
  - Better a false positive than a false negative

Case study: the 2014 Genoa Flooding

Total annotations:

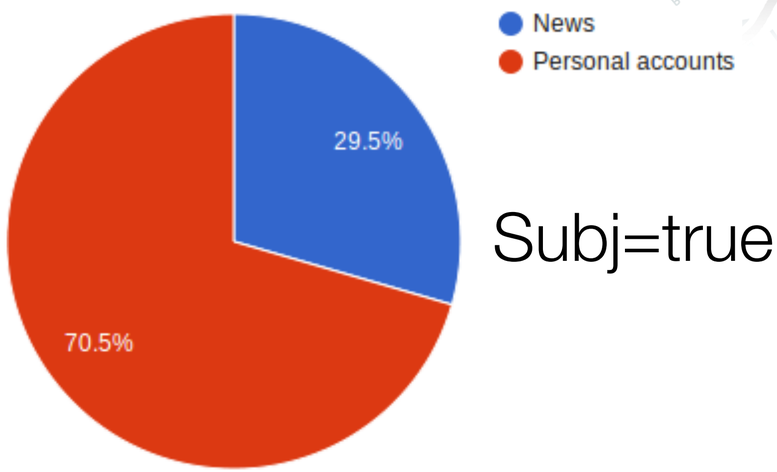
8922 subj
499 pos
3519 neg
1019 iro

6033 users

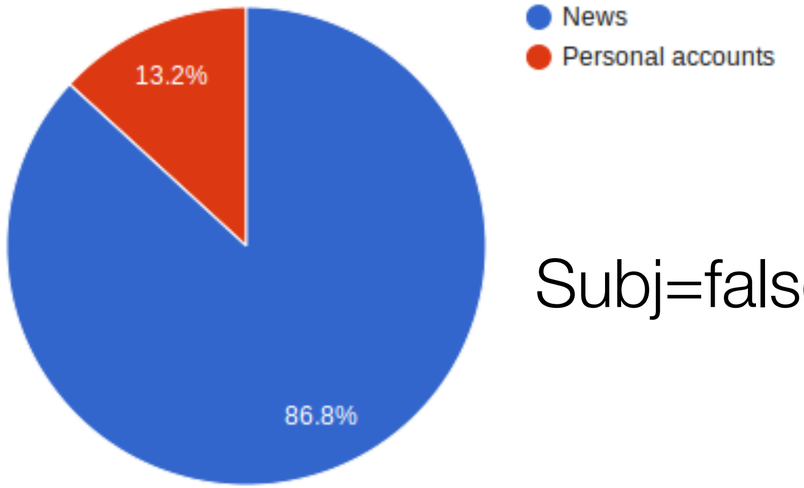
Problem: evaluation

Evaluation of the Accuracy in Subjectivity Detection

- Manual annotation of 75 sources
- **News** (59.2%)
  - @ScoopSquareGE, @MeteoWeb_eu, @infoitinterno, @ITnewsGE, @TopTrendIT…
- **Personal accounts** (40.8%)
  - @Miti_Vigliero, @mauneobux, @LaRouge_DOC, …



Subj=true



Subj=false

(Indirect) evaluation - relevant items

- Subset of relevant hashtags, toponyms and topics :
  - #allertameteo, #protezionecivile, #alluvionege, ...
  - Montoggio, Sturla, Fereggiano, ...
  - “Ondata di piena”, “Invaso dal fango”, “esonda …”…

(Indirect) evaluation - relevant items

*acqua , deraglia un freccia
*allagamenti e paura
*allerta massima
*allerta meteo
*alluvione in corso
*croce rossa
*crolla muraglione
*danni enormi
*danni ingenti
*deraglia un treno
*emanata allerta
*esce dai binari
*esonda il bisagno
*esonda lo scrivia
*esondano i fiumi
*esondano il bisagno
*esondano i torrenti
*esonda un torrente
*evacuate 16 famiglie
*fa paura
*frana improvvisa
*ho paura
*invaso dal fango

* livello di allerta
* mancato allarme
* meteo da allerta
* momenti di ansia
* morto e dispersi
* morto e gravi
* morto foto
* nuova alluvione
* nuovo alluvione
* ondata di piena
* piani alti
* prestare massima attenzione
* protezione civile
* rubano nei negozi
* sciacalli in azione
* situazione critica
* situazione drammatica
* tanta paura
* trascinate da acqua
* trascinate dall' acqua
* travolge auto
* trovato il corpo
* usate l' auto
* uscire di casa
* uscite da casa

* #allerta2
* #allertameteo
* #allertameteoge
* #allertameteogenova
* #allertameteolg
* #allertameteolig
* #alluvione
* #alluvionege
* #angelidelfango
* #arpal
* #bisagno
* #busalla
* #campoligure
* #chiaravagna
* #crocerossa
* #emergenza24
* #emergenze

#entella
#fereggiano
#forzagenovarisorgidalfango
#genovaalluvione
#iononrischio
#maltempo
#molassana
#montoggio
#noncefangochetenga
#ovada
#polcevera
#protezionecivile
#stura
#voltri

Detecting trending items with the Poisson law

- Extraction of trending* (hashtags | toponymes | topics) for each hour

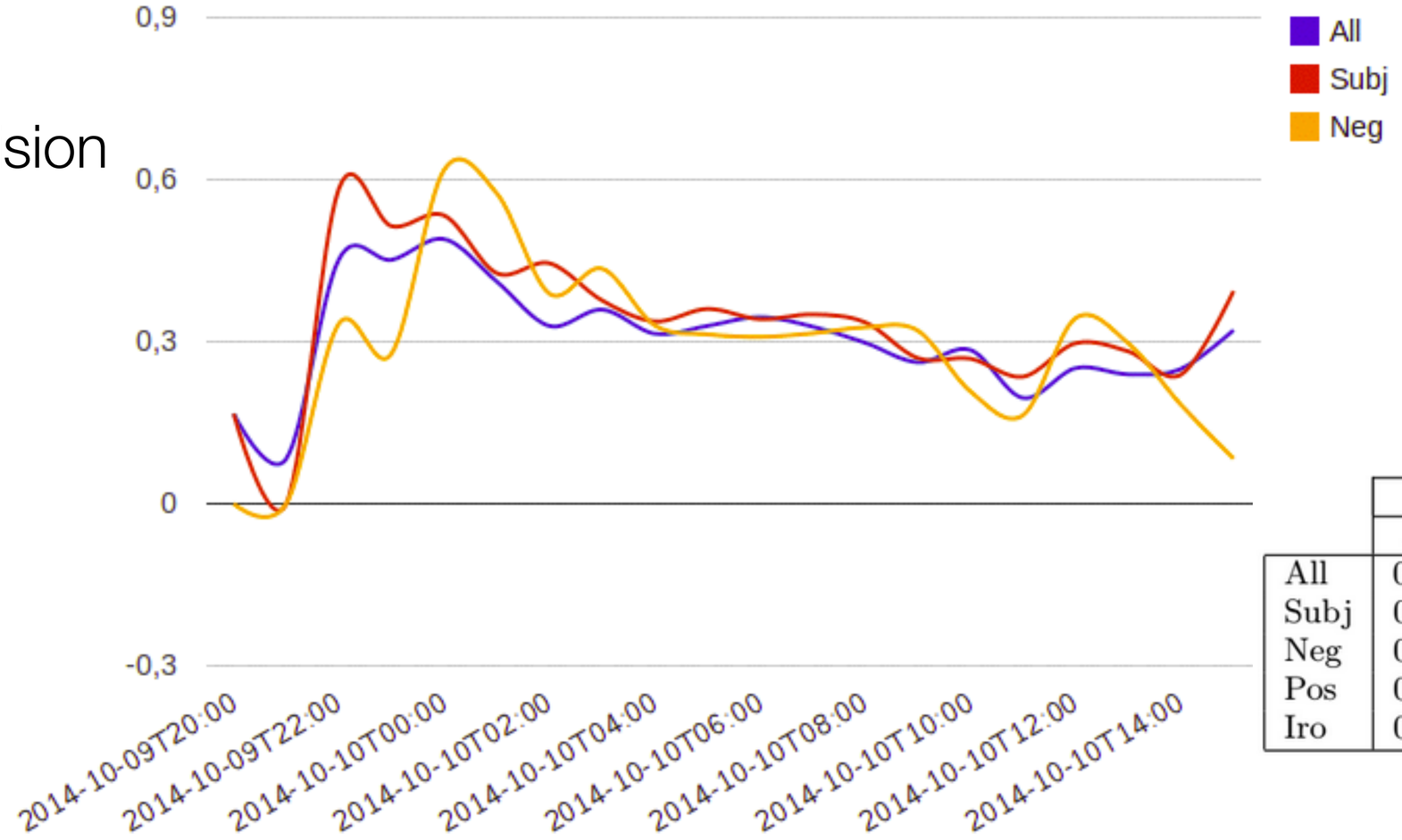
$$Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} > \epsilon$$

- The key is to model λ to capture the expected frequency of an item
  - With enough data, it is possible to model the frequency over the same period (week, hour, etc.) of the year
- In our test, the threshold ε was set to 0
- λ was set to the estimated frequency during the previous hour

(Indirect) evaluation

- plot precision (# relevant trending / # trending)

Precision



	Hashtags		Toponyms		Topics	
	Acc.	Cov.	Acc.	Cov.	Acc.	Cov.
All	0.160	1.000	0.660	0.875	0.104	0.980
Subj	0.197	0.875	0.661	0.844	0.156	0.804
Neg	0.254	0.625	0.482	0.594	0.137	0.451
Pos	0.208	0.188	0.155	0.156	0.164	0.157
Iro	0.322	0.313	0.362	0.344	0.118	0.098

Mean Average Precision

Qualitative analysis

- Most of the **positive** tweets (82 out of 493 - ~16%) came from the same account (*spam?*):
  - "#News e #Astronomia Genova alluvionata, esondano i fiumi: un morto... -... http://bit.ly/110yfvW *Buona visione ! :))*"
- Many false positives for **irony** but also some good answers: "Ragazzi tranquilli #Renzi ha detto che non ci lascerà soli"
- **Negative** tweets may convey different emotions:
  - Fear: "Basta guardare fuori per avere paura"
  - Rage: "In tre anni non è stato fatto nulla, amministrazioni delinquenti"
  - Sadness: "Mi viene da piangere"
- Some features may change polarity (meaning) depending on the **context**:
  - "Coraggio Genova!!!" vs. "Salite ai piani alti!!!"
  - "Speriamo bene" vs. "Stiamo bene"

What we did learn

- We can assume that the positive tweets in such a context could be considered **spam**
- Need to distinguish between **different emotion types**:
  - Worries, indignation, criticism, rage, etc.
- The set of subjective tweets may provide some insights about a critical situation faster than looking at the whole set of tweets
- Irony: still far from having usable results, but it may be useful to discover criticism and unconstructive opinions
- Need to improve performances and build a dedicated evaluation framework

Conclusions

Sentiment Analysis: where are we?

- Best system at SemEval-2015 was a mix of 3 pre-existing systems

#	System	2013: Progress		2014: Progress			2015: Official	
		Tweet	SMS	Tweet	Tweet sarcasm	Live-Journal	Tweet	Tweet sarcasm
1	Webis	68.49_{10}	63.92_{14}	70.86_{7}	49.33_{12}	71.64_{14}	$\mathbf{64.84}_{1}$	53.59_{22}
2	unitn	72.79_{2}	68.37_{2}	73.60_{2}	55.44_{5}	72.48_{12}	$\mathbf{64.59}_{2}$	55.01_{19}
3	lsislif	71.34_{4}	63.42_{17}	71.54_{5}	46.57_{22}	73.01_{10}	$\mathbf{64.27}_{3}$	46.00_{33}
4	INESC-ID*	71.97_{3}	63.78_{15}	72.52_{3}	56.23_{3}	69.78_{22}	$\mathbf{64.17}_{4}$	64.91_{2}
5	Splusplus	72.80_{1}	67.16_{5}	74.42_{1}	42.86_{31}	75.34_{1}	$\mathbf{63.73}_{5}$	60.99_{7}

- Performances are still low, even in a binary classification scenario
- Irony/sarcasm constitute a hard problem
- Testing limits of existing NLP techniques?

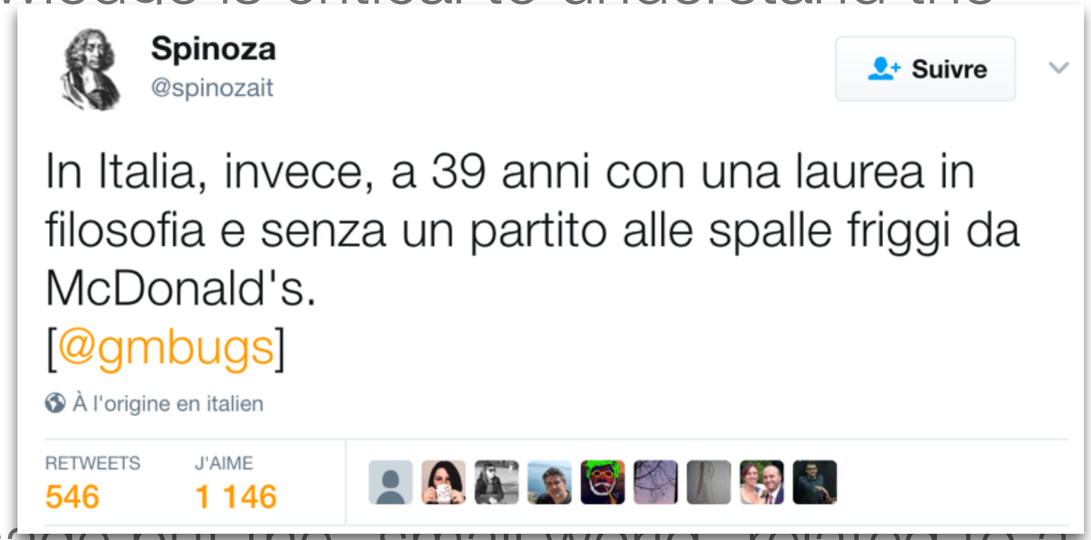
Sentiment Analysis: where are we?

- Dataset quality problem
  - Example: SENTIPOLC:
- Detecting irony or just thematic and/or stylistic variations?
- Subjectivity: are we building a scanner for http addresses?

Word	IRO	p(w|iro)	NOT IRO	p(w|notiro)	diff prob
Mario	139	0,016048955	490	0,009626719	0,006422236
Monti	261	0,030135088	1218	0,023929273	0,006205815
governo	166	0,019166378	738	0,014499018	0,00466736
che	155	0,017896317	717	0,014086444	0,003809873
?	169	0,019512758	865	0,016994106	0,002518652
!	182	0,02101374	970	0,019056975	0,001956765
#Monti	72	0,008313128	327	0,006424362	0,001888766
)	61	0,007043067	281	0,005520629	0,001522438
Italia	12	0,001385521	167	0,003280943	
anche	19	0,002193742	216	0,004243615	
non	88	0,01016049	637	0,012514735	
della	28	0,003232883	287	0,005638507	
via	6	0,000692761	164	0,003222004	
fb	8	0,000923681	219	0,004302554	
0	10	0,001154601	253	0,004970531	
bit	13	0,001500981	305	0,005992142	
//www	12	0,001385521	332	0,006522593	
…	24	0,002771043	413	0,008113949	
#Grillo	26	0,003001963	591	0,011611002	
//	39	0,004502944	920	0,018074656	
http	53	0,006119386	1239	0,024341847	
:	219	0,025285764	2460	0,048330059	

Sentiment Analysis: where are we?

- Few works go “outside” the training corpora
  - In most cases (if not all), external knowledge is critical to understand the polarity of the message
  - Who is the emitter of the message?
  - What is referenced by the message?



- In any case, we need not only the message but the “small world” related to a given message

References

- Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.
- NLP Course by Dan Jurafsky : https://web.stanford.edu/~jurafsky/NLPCourseraSlides.html
- Buscaldi D., Hernández-Farias I. Sentiment Analysis on Microblogs for Natural Disasters Management: a Study on the 2014 Genoa Floodings. In: Proceedings of the 3rd International Workshop on Social Web for Disaster Management (SWDM'15), co-located with the WWW'15 conference. pp.1185-1188. Firenze, Italia.
- Hernández-Farias I., Buscaldi D., Priego-Sánchez, B. IRADABE: Adapting English Lexicons to the Italian Sentiment Polarity Classification task. In: Proceedings of the First Italian Conference on Computational Linguistics (CLiC-it 2014) and the fourth International Workshop EVALITA2014, pp.75-81. Pisa, Italia.