SOCIAL MEDIA ANALYTICS

Sentiment Analysis

INTRODUCTION

INTRODUCTION

IN 1 SECOND…

54,977 Facebook posts

1,222 Tumblr posts

7,553 Tweets sent

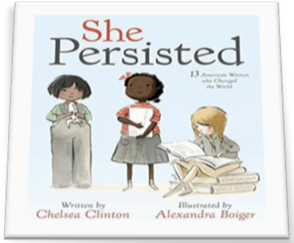
350,000 tweets sent per **minute**

500 million tweets per **day**

200 billion tweets per **year**.

WHY SOCIAL MEDIA

- Fresh information emerging in real-time on microblogs:
  - **New** (relevant/popular) **entities** (Book Launch)



  - **New events** (London Shooting)

  - **Factual information** (Death of Chuck Berry)

  - **New relations** (Donald Trump: POTUS)

INFORMATION VS KNOWLEDGE

- **Information** is a collection of **facts**
  - Reports about an event or activity

- **Knowledge**
  – Awareness and understanding of a set of information and the ways that information can be made useful
  – Derived and/or inferred from raw information in order to enable informed decisions

SOCIAL MEDIA ANALYTICS

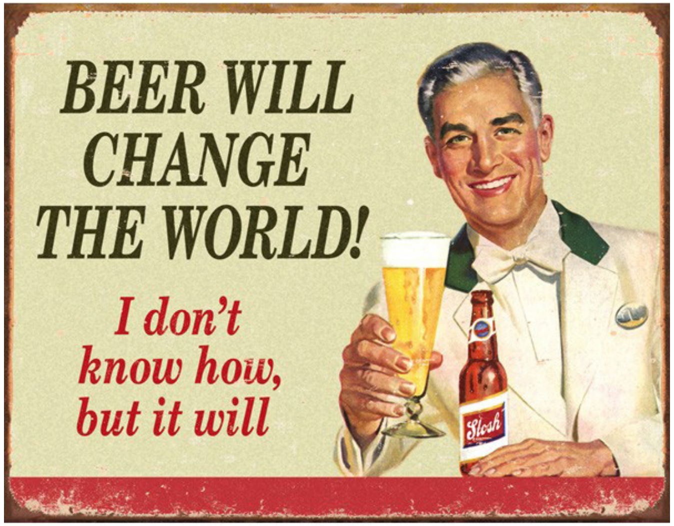
We are in the era of Social Media Analytics, which is the *non-trivial* process of identifying

- *valid*
- *novel*
- potentially *useful* information
- to be transformed into ***actionable knowledge*** from social media

When dealing with social media text we need to make sense and transform them into **actionable knowledge**

WHY SOCIAL MEDIA ANALYTICS

Brand Reputation



Quality of life

Predicting market data

WHY SOCIAL MEDIA ANALYTICS

- **Opinions** given by others have an impat on our choices

20 K

1.2 millions

36 millions

290 millions

41 K/s

278 K/s

Social platforms: source of **huge volumes of data**

WHY SOCIAL MEDIA ANALYTICS

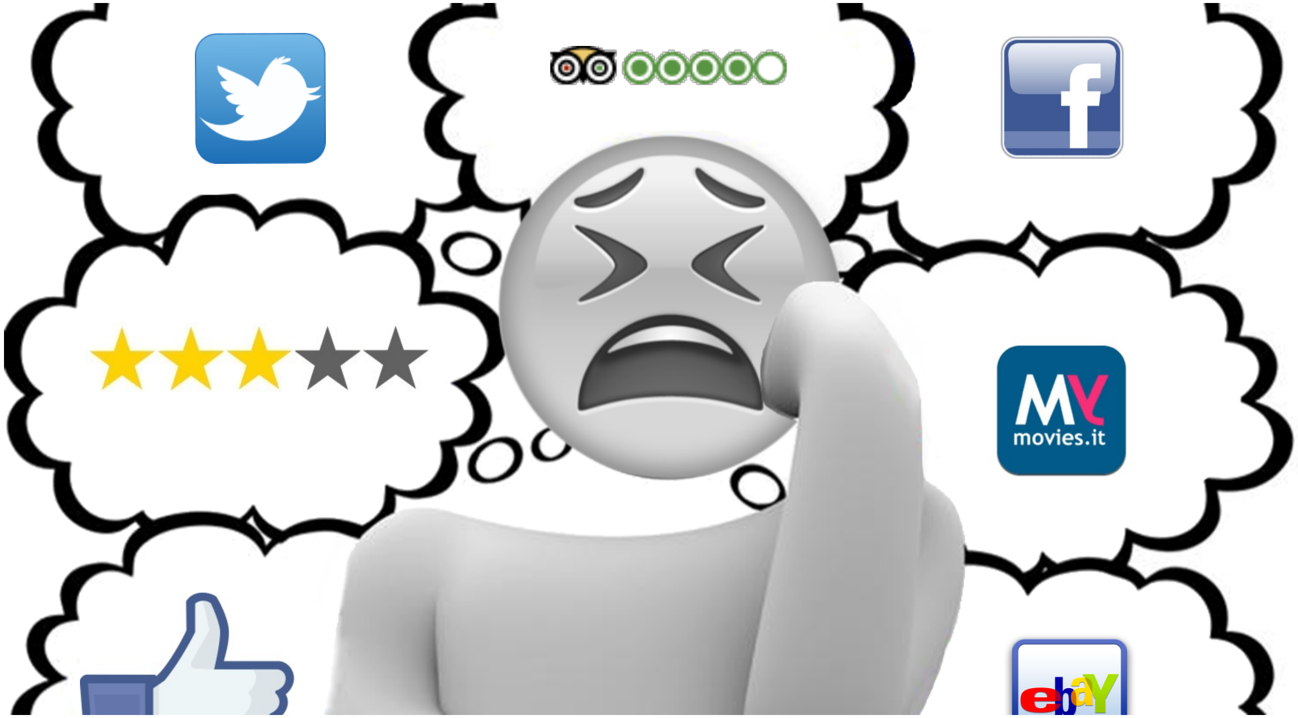
Now we have "a lot" of informamation about a given argument

… we can analyze them to make our decisions ….

WHY SOCIAL MEDIA ANALYTICS

Easy?

WHY SOCIAL MEDIA ANALYTICS



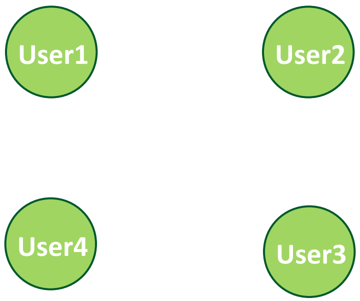
Too much information

=

No information

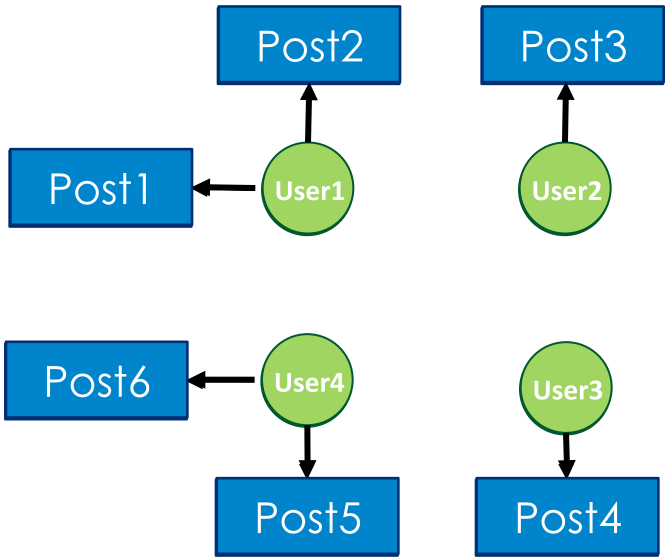
BASIC COMPONENTS

- **Basic components** for SMA:
  - **Opinion holder (user)**: The person or organization that holds a specific opinion on a particular object.



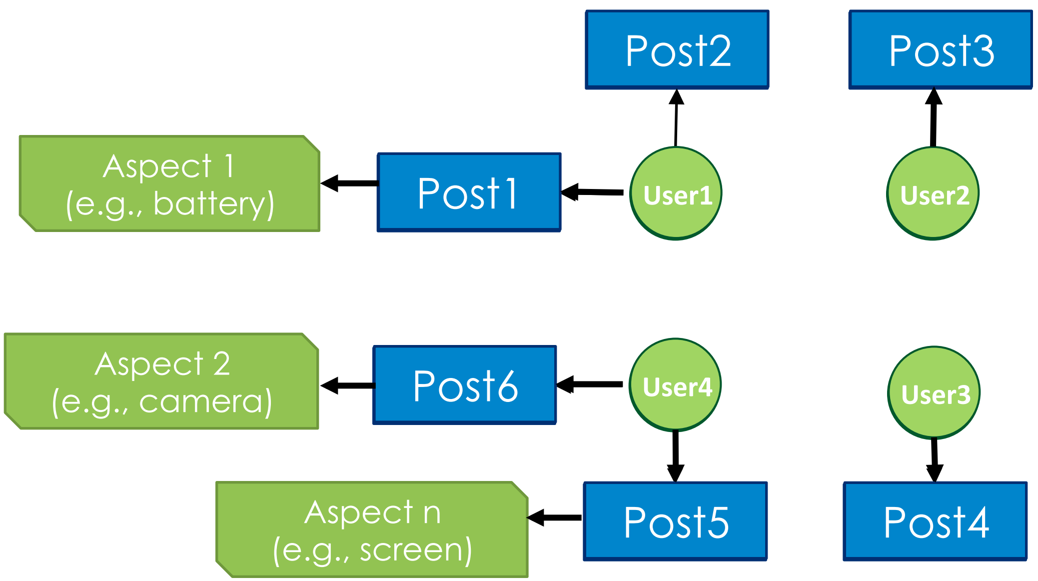
BASIC COMPONENTS

- Basic components for SMA:
  - **Opinion holder (user)**: The person or organization that holds a specific opinion on a particular object.
  - **Object**: on which an opinion is expressed in a post



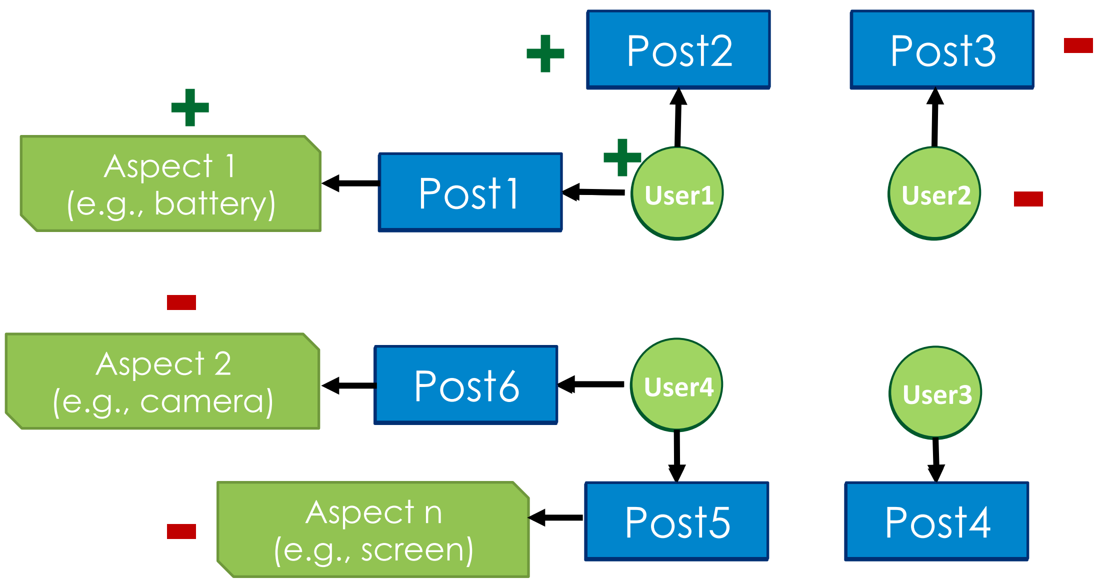
BASIC COMPONENTS

- Basic components for SMA:
  - **Opinion holder (user)**: The person or organization that holds a specific opinion on a particular object.
  - **Object**: on which an opinion is expressed.
  - **Aspect**: aspect of the object.



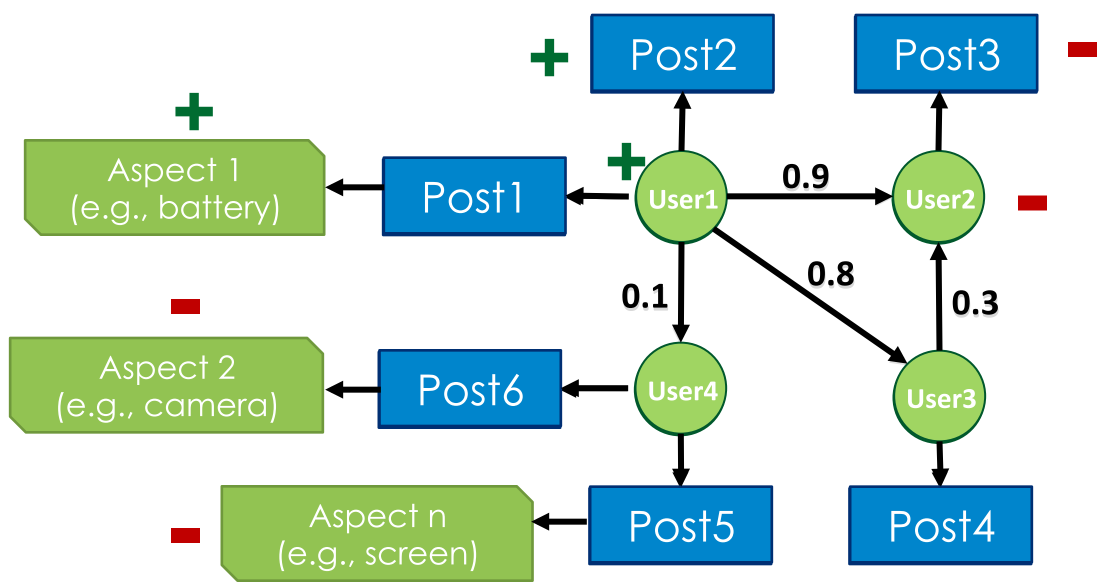
BASIC COMPONENTS

- **Basic components** for SMA:
  - **Opinion holder (user)**: The person or organization that holds a specific opinion on a particular object.
  - **Object**: on which an opinion is expressed.
  - **Aspect**: aspect of the object.
  - **Opinion**: a view, attitude, or appraisal on an object from an opinion holder.



BASIC COMPONENTS

- **Basic components** for SMA:
  - **Opinion holder (user)**: The person or organization that holds a specific opinion on a particular object.
  - **Object**: on which an opinion is expressed.
  - **Aspect**: aspect of the object.
  - **Opinion**: a view, attitude, or appraisal on an object from an opinion holder.
  - **(Social Network)**



WHY IS IT SO DIFFICULT?

REPRESENTATION

WHY IS IT SO DIFFICULT?

Get the cat with the gloves.

AMBIGUITY

AMBIGUITY

- Ambiguity is a crucial problem for natural language understanding/processing.

Paris was shot in the hand as she chased

the robbers in the back street

AMBIGUITY

- Morpho-syntactic ambiguity:
  - **PART-OF-SPEECH TAGGING**

Paris *was* ***shot*** *in the* ***hand*** *as she* ***chased***

…e … s in t… k stree…

NN VB

NN VB

NN VB

JJ VB

AMBIGUITY

- Morpho-syntactic ambiguity:
  - **PART-OF-SPEECH TAGGING**

Paris *was* ***shot*** *in the* ***hand*** *as she* ***chased***

e *s in t* *k stree*

NN ~~VB~~

~~NN~~ **VB**

NN ~~VB~~

~~JJ~~ **VB**

AMBIGUITY

- Structural (syntactic) ambiguity:
  - **PARSE TREE DISAMBIGUATION**

Paris was shot in the hand as she chased

the robbers ***in the back street***

AMBIGUITY

- Structural (syntactic) ambiguity:
  - **PARSE TREE DISAMBIGUATION**

Paris was shot in the hand as she chased

the robbers ***in the back street***

AMBIGUITY

- Structural (syntactic) ambiguity:
  - **PARSE TREE DISAMBIGUATION**

Paris was shot in the hand as she ***(chased***
(the robbers)$_{NP}$ ***(in the back street)$_{PP}$)***

AMBIGUITY

- Semantic ambiguity:
  NAMED-ENTITY RECOGNITION
  NAMED-ENTITY LINKING

New Frozen Boutique to Open at Disney's Hollywood Studios. Can't wait 😢

dia.org/page/The_Walt_Disney_Company

http://dbpedia.org/page/Frozen_(2013_film)

https://en.wikipedia.org/wiki/Frozen_(Madonna_song)

http://dbpedia.org/page/Disney's_Hollywood_Studios

HIDDEN SEMANTICS

- Emotional ambiguity:

SENTIMENT ANALYSIS

IRONY DETECTION

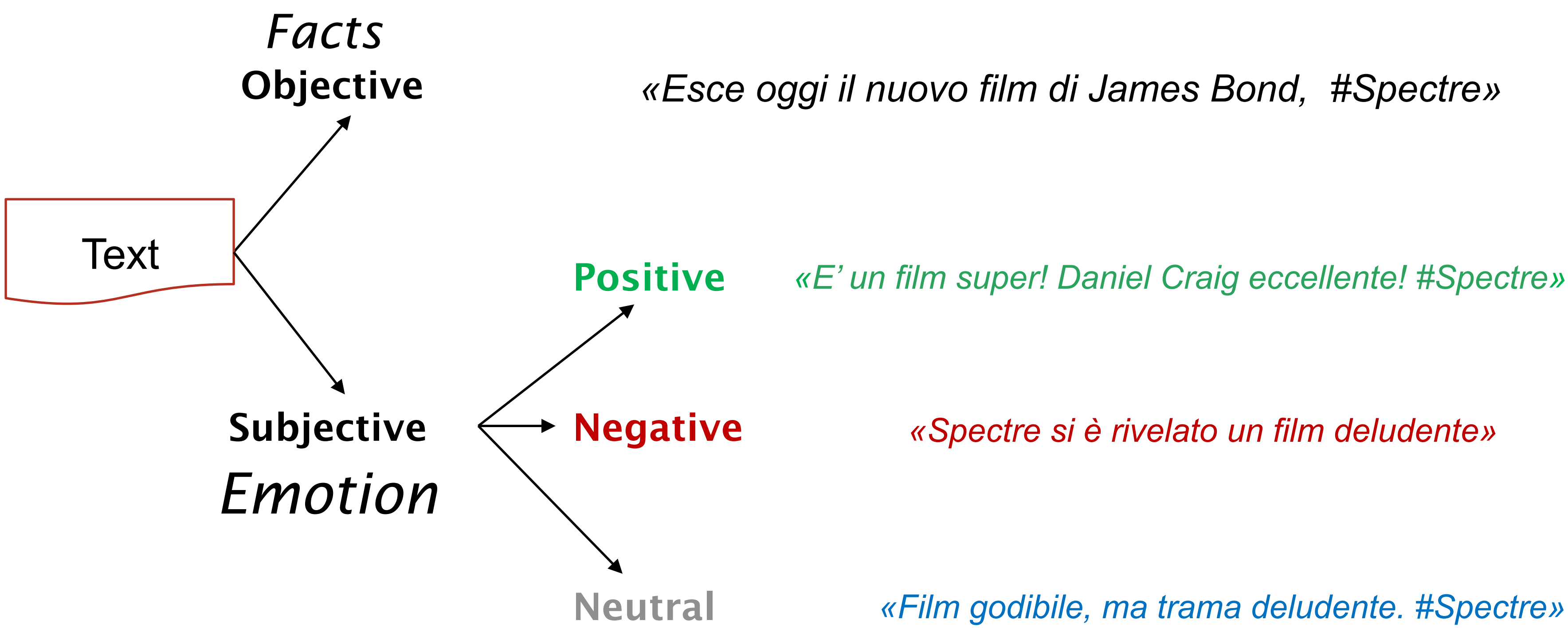
New Frozen Boutique to Open at Disney's Hollywood Studios. Can't wait 😢

SENTIMENT ANALYSIS

Subjectivity, Polarity and Irony Detection

INTRODUCTION

Social media analytics should be able to **interpret** natural language to identify **subjective** information that denote opinions and sentiments, determine the corresponding **polarity** (positive, negative, neutral) e understand the targeted **subject/object**.



WHERE CAN WE FIND OPINIONS?

Customer Reviews

281

4.8 out of 5 stars

5 star		86%
4 star		7%
3 star		3%
2 star		1%
1 star		3%

Nikon D5500 DX-format Digital SLR Dual Lens Kit w/ - Nikon AF-...

by Nikon

Color: Black | Style: Dual Lens Kit (18-55mm VR & 70-300mm Lens) | Configuration: Base | Change

Price: $596.95 + Free shipping with Amazon Prime

Rate this item

Write a review

Top positive review

See all 262 positive reviews ›

315 people found this helpful

I decided to gear down and sell all of my Nikon equipment and buy a good point and shoot

By Ed Hays on July 19, 2015

My situation may be a little different than some others on this forum. My first DSLR was a Nikon D90 and then I moved to a D7000. I had assorted Nikon lenses and equipment and having reached age 70, I decided to gear down and sell all of my Nikon equipment and buy a good point and shoot. After looking into everything that was available in the point n shoot world, I realized that nothing would take the place of of my D7000; that is until the D5500 was introduced. Between its amazingly light weight, articulating touch screen LCD, WiFi, uniformly

Read more

Top critical review

See all 19 critical reviews ›

8 people found this helpful

Very disappointed

By Amazon Customer on December 1, 2016

I had a D40, which eventually stopped working. I replaced it with a D5100, which also after only 2 years started giving me error messages when I pushed the shutter button. I have an expensive Nikon lens or I would not have bought another Nikon. I bought the D5500 two weeks before a big trip to Morocco. I tested it before I left & it worked okay. I got to Morocco & about every 7th or 8th time I pushed the shutter button, I got no response. It would not even let me review previous taken pictures. I had to keep shutting off & on the camera sometimes numerous times before the shutter fired. I am not ready to deal with

Read more

SOME BASIC DEFINITIONS

- Definitions:

  An ***objective*** post p_0 presents some factual information about the world, while a ***subjective*** post p_s expresses some personal feelings, views, or beliefs. [Liu 2012]

[subjective] “I’ve just watched #Deadpool… I LOVE IT!!!! ”

[objective] “Tomorrow #SuicideSquade with my baby”

SOME BASIC DEFINITIONS

- Definitions: Positive and Negative are "simple"

  A ***neutral*** post is a message that lies between positive and negative [Pang and Lee 2008]

[positive] "*Best Joker EVER!! #suicidesquad*"

[negative] "*Deadpool is so childish! I slept during the movie*"

[neutral] "*Goodmovie,@VancityReynolds worst actor ever #deadpool*"

SOME BASIC DEFINITIONS

- Definitions:

An ***explicit*** *opinion* is a subjective statement that gives an opinion. [Liu 2012]

An ***implicit*** *opinion* is an objective statement that implies an opinion. [Liu 2012]

[explicit] "*Suicide Squad is a great movie with an awesome cast!*"

[implicit] "*I went out the cinema after 15 minutes #suicidesquad*"

SOME BASIC DEFINITIONS

- Definitions:

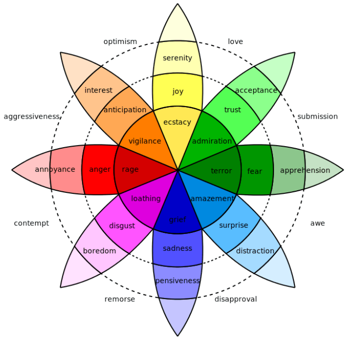
  An ***ironic*** *post* is a communicative way that expresses the opposite of what is literally said. [Wilson and Sperber 2007]

  [ironic] "*Suicide Squad Very nice movie...#irony*"

SOME BASIC DEFINITIONS

- Definitions:

    Emotions. [Plutchik 2001]



[Anger] *"#Deadpooll wasted time and money grrrrrrrr"*

[Anticipation] *"Can'twait to see Deadpool!!!"*

[Joy] *"Deadpool was A-M-A-Z-I-N-G"*

[Trust] *"Best movie ever #Deadpool! Trust me!"*

[Fear] *"Saw #Deadpool last night. I was frightened during some crude scenes"*

[Surprise] *"Much to my surprise, I actually liked Deadpool."*

[Sadness] *"I finally got to watch deadpool and im so sad this is so boring"*

[Disgust] *"Deadpool is everything I hate about our century combined in the trashiest movie possible."*

SMA PROCESS

Collect

Represent

Classify

Predict

Decide

COLLECT

- Real time expressions (on-line)

  “What’s happening now?”

- Messaggi scaricati off-line

  “What do people think about iPhone7?”

REPRESENT

- How can we represent social media text?
  - It's necessary to move from **qualitative data** (text) to **quantitative data (numbers)**…to **measure, compare and learn**

«E' un film super! Daniel Craig eccellente!#Spectre »

« Spectre si è rivelato un film deludente »

«Un film godibile, ma trama deludente. #Spectre »

film	1	1	1
super	1	0	0
eccellente	1	0	0
spectre	1	1	1
rivelato	0	1	0
deludente	0	1	1
godibile	0	0	1
trama	0	0	1

REPRESENT

Preprocessing techniques:

1. Stopwords removal(articles, conjunctions, …)
2. Numbers and punctuation removal (not always true...)
3. Stemming (es: ‘*gatto*’, ‘*gatta*’, ‘*gattino*’, ‘*gattaccio*’ -> ‘*gatt*’)

REPRESENT

Social networks text can contain other **important characteristics**:

1. Word lenghtened (es: beeeeeeellllooooo)
2. Onomatopoeic expressions (es: bleh)
3. Slang (es: grz-> grazie)
4. Acronyms (es: ROFL)
5. Emoji (es: 😂, 😊, 😭)

Onomatopoeic Lenghtening *Slang* Acronym

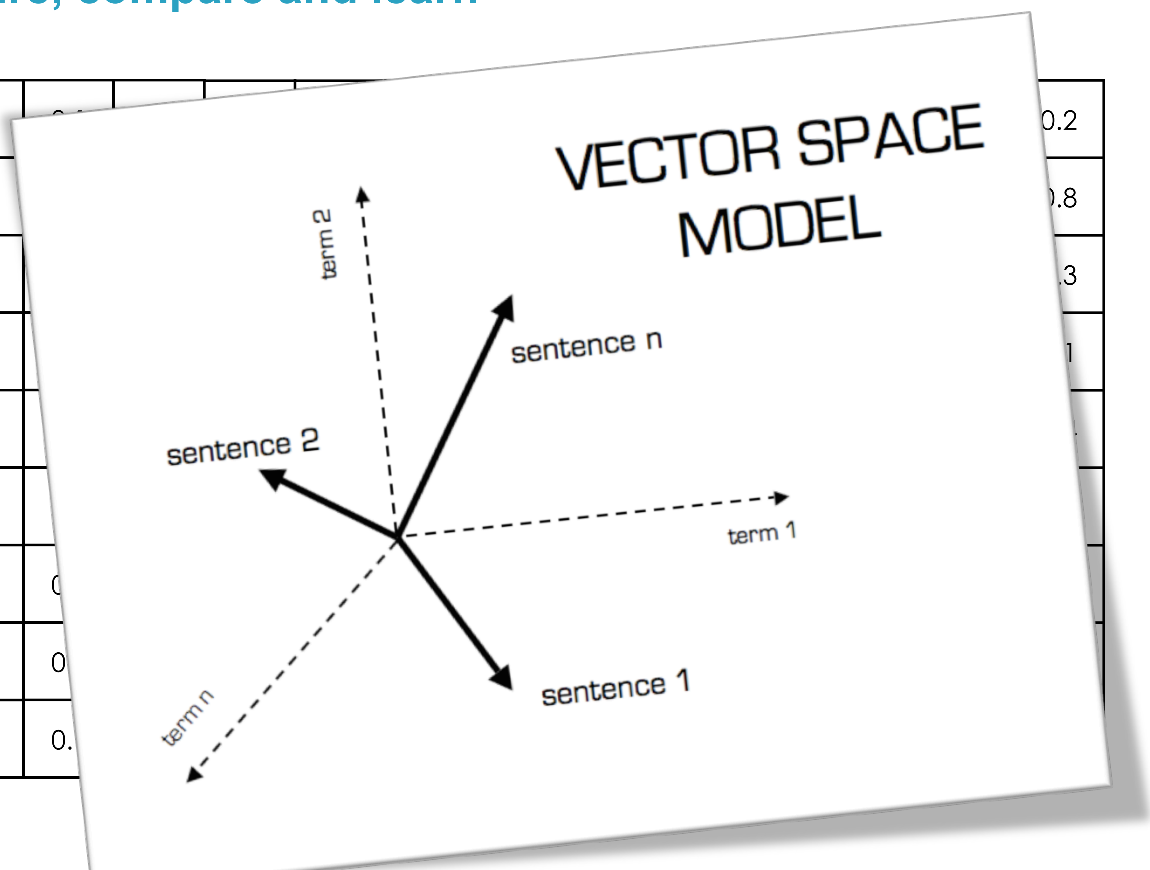
Wow its beautifuuuuulll!! Tks god iPhone 5S is really cool!!!! 😊 ROFL

Emoji

REPRESENT

- How can we represent social media text?
  - It's necessary to move from **qualitative data** (text) to **quantitative data (numbers)**…to **measure, compare and learn**

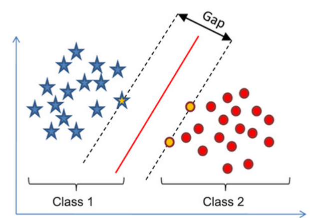
film	0.2	0.1	0.3	0.2
super	0.8	0.2	0.9	0.8
eccellente	0.3	0.9	0.2	0.3
spectre	0.1	0.7	0.2	0.1
rivelato	0.4	0.6	0.3	0.4
…	…	…	…	…
Slang	0.7	0.2	0.6	0.7
Acronimi	0.6	0.1	0.5	0.6
😊	0.9	0.5	0.6	0.9



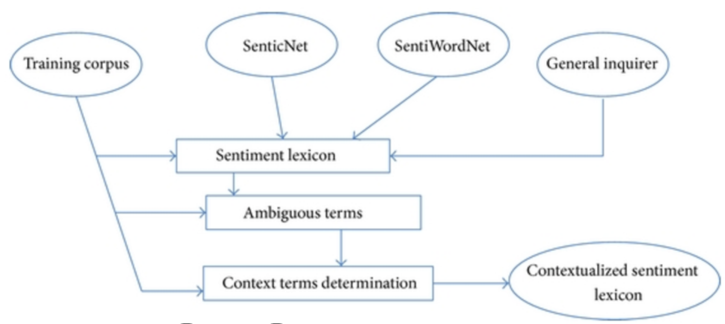
CLASSIFY

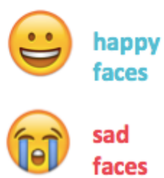
- Lexicon-based approach

- Supervised learning

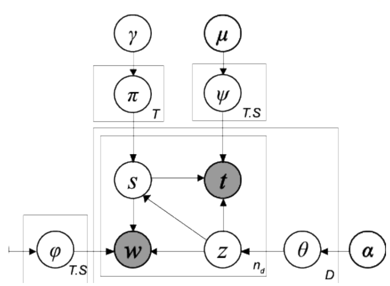


- Semi-Supervised learnin





- Unsupervised learning



LEXICON-based APPROACHED

«E' un film super ! Daniel Craig eccellente !#Spectre »

film	1	1	1
super	1	0	0
eccellente	1	0	0
spectre	1	1	1
rivelato	0	1	0
deludente	0	1	1
godibile	0	0	1
trama	0	0	1

LEXICON-based APPROACHED

«Spectre si è rivelato un film deludente »

film	1	1	1
super	1	0	0
eccellente	1	0	0
spectre	1	1	1
rivelato	0	1	0
deludente	0	1	1
godibile	0	0	1
trama	0	0	1

LEXICON-based APPROACHED

«E' un film godibile , ma
trama deludente.#Spectre »

film	1	1	1
super	1	0	0
eccellente	1	0	0
spectre	1	1	1
rivelato	0	1	0
deludente	0	1	1
godibile	0	0	1
trama	0	0	1

SOME LEXICON

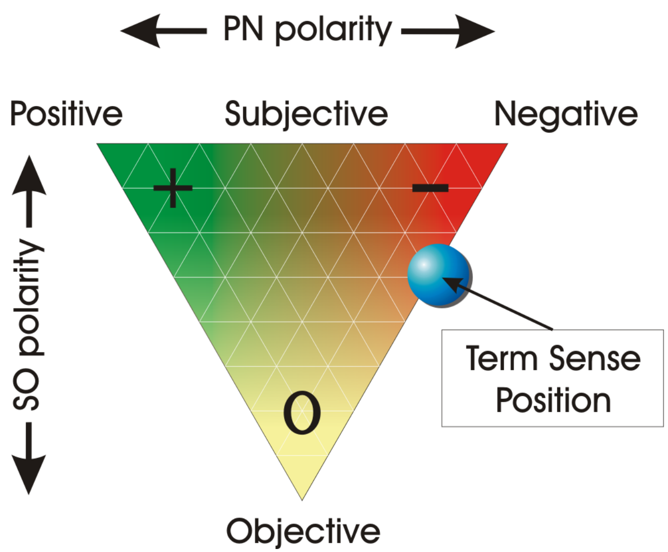
- [Bing Liu's Page on Opinion Mining]
- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

- 6786 words
  - 2006 positive
  - 4783 negative

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

SENTIWORDNET

- Home page: http://sentiwordnet.isti.cnr.it/
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness



Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010 SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC-2010

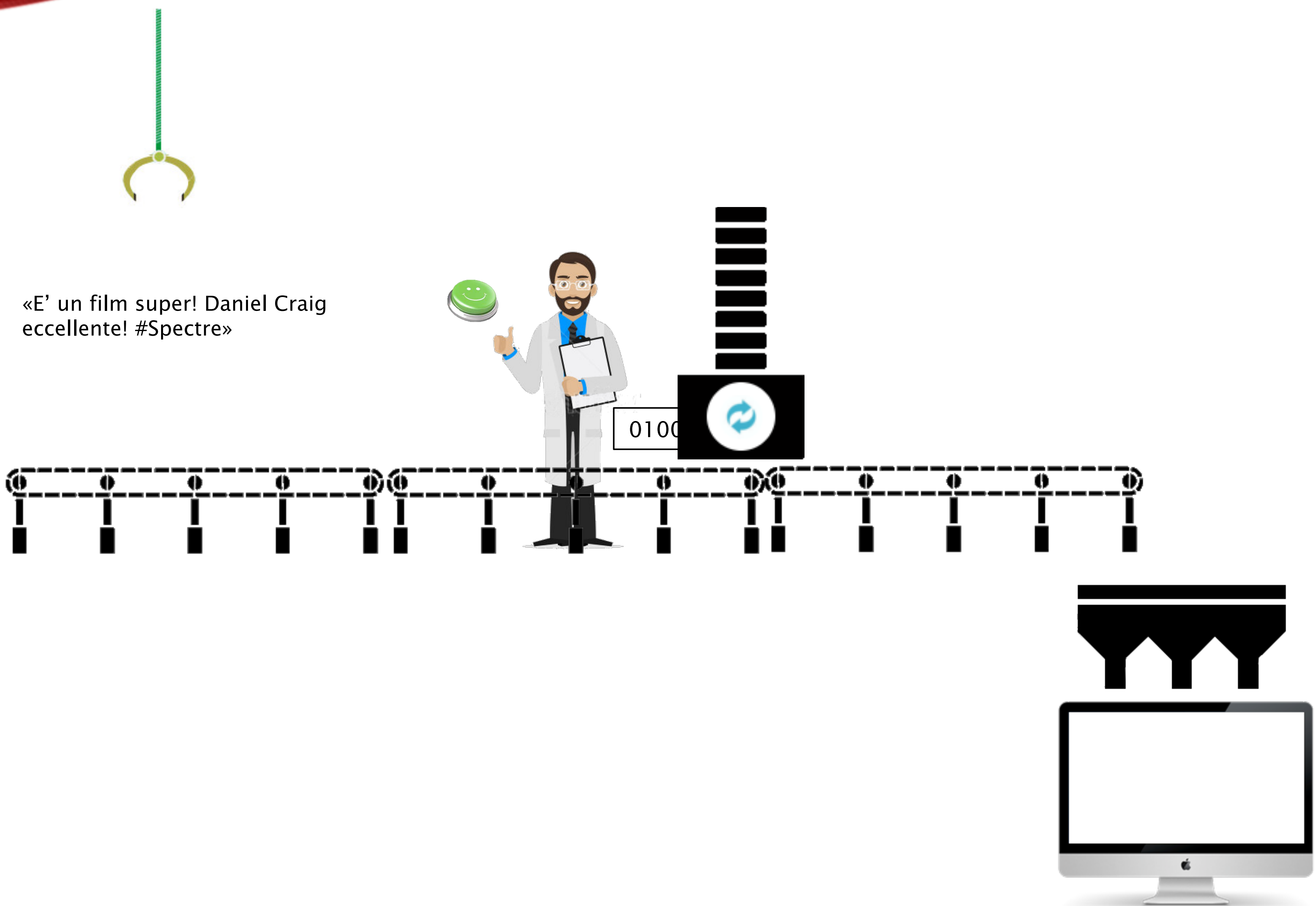
SENTIWORDNET

- Sentiment word decision:
  - computing the number of times the ‘word#sense’ entry is more positive than negative (positive **>** negative) and vice-versa in SentiWordNet.
- If positive > negative, then the word is positive
- If negative> positive, then the word is negative
- If there is no variation between positive and negative, then the word in neutral.

(SOME) OTHER LEXICONS

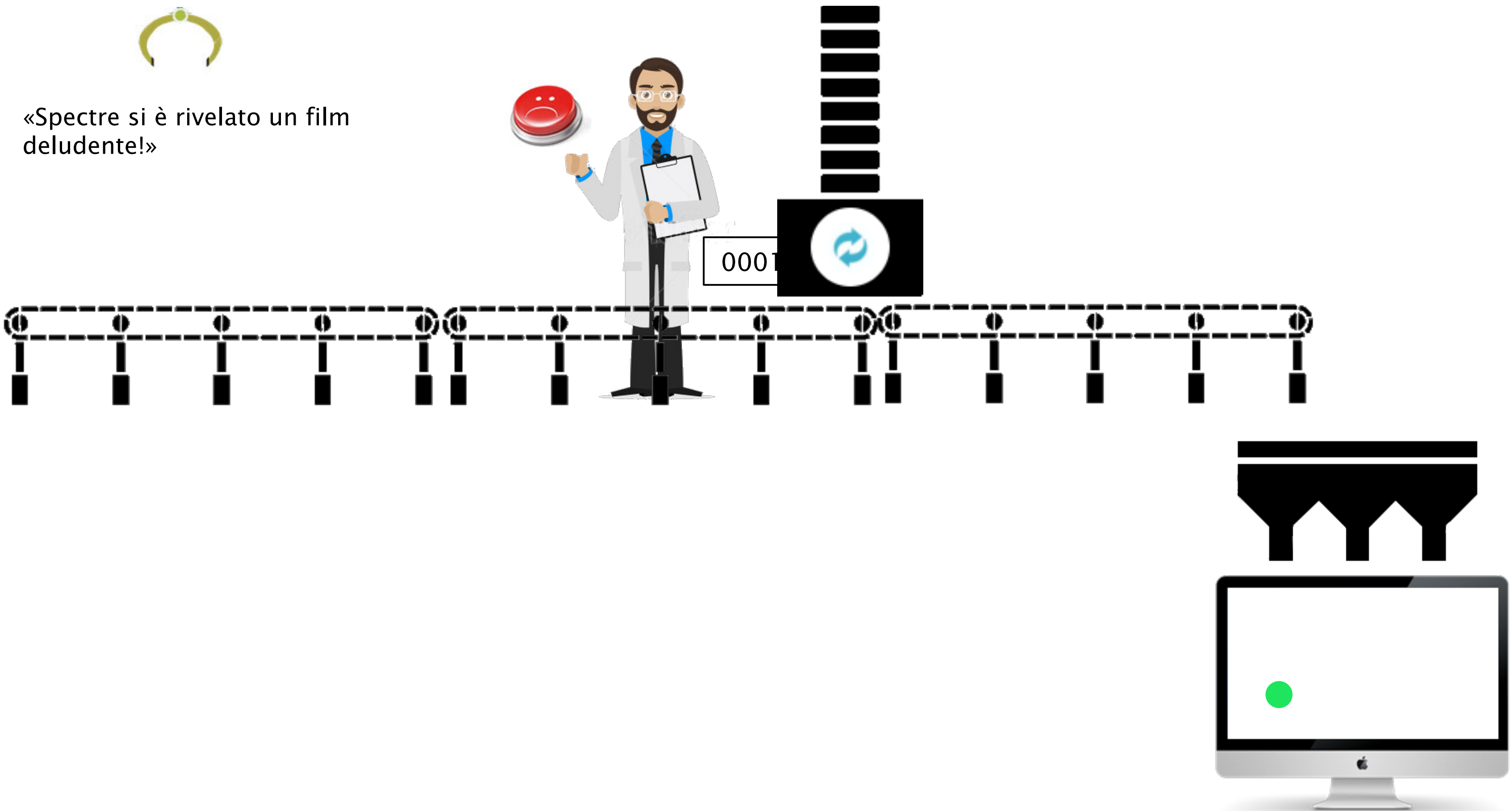
- The Subjectivity Sense Annotation
  - http://mpqa.cs.pitt.edu/lexicons/subj_sense_annotations/
- Yelp Lexicon (about restaurants)
  - http://www.yelp.com/dataset_challenge
- AFINN (-5;+5)
  - https://github.com/abromberg/sentiment_analysis/blob/master/AFINN/AFINN-111
- Sentiment140 (twitter lexicon)
  - http://saifmohammad.com/WebPages/lexicons.html
- Sentix (Italian language)
  - http://www.let.rug.nl/basile/twita/downloads.php

SUPERVISED LEARNING

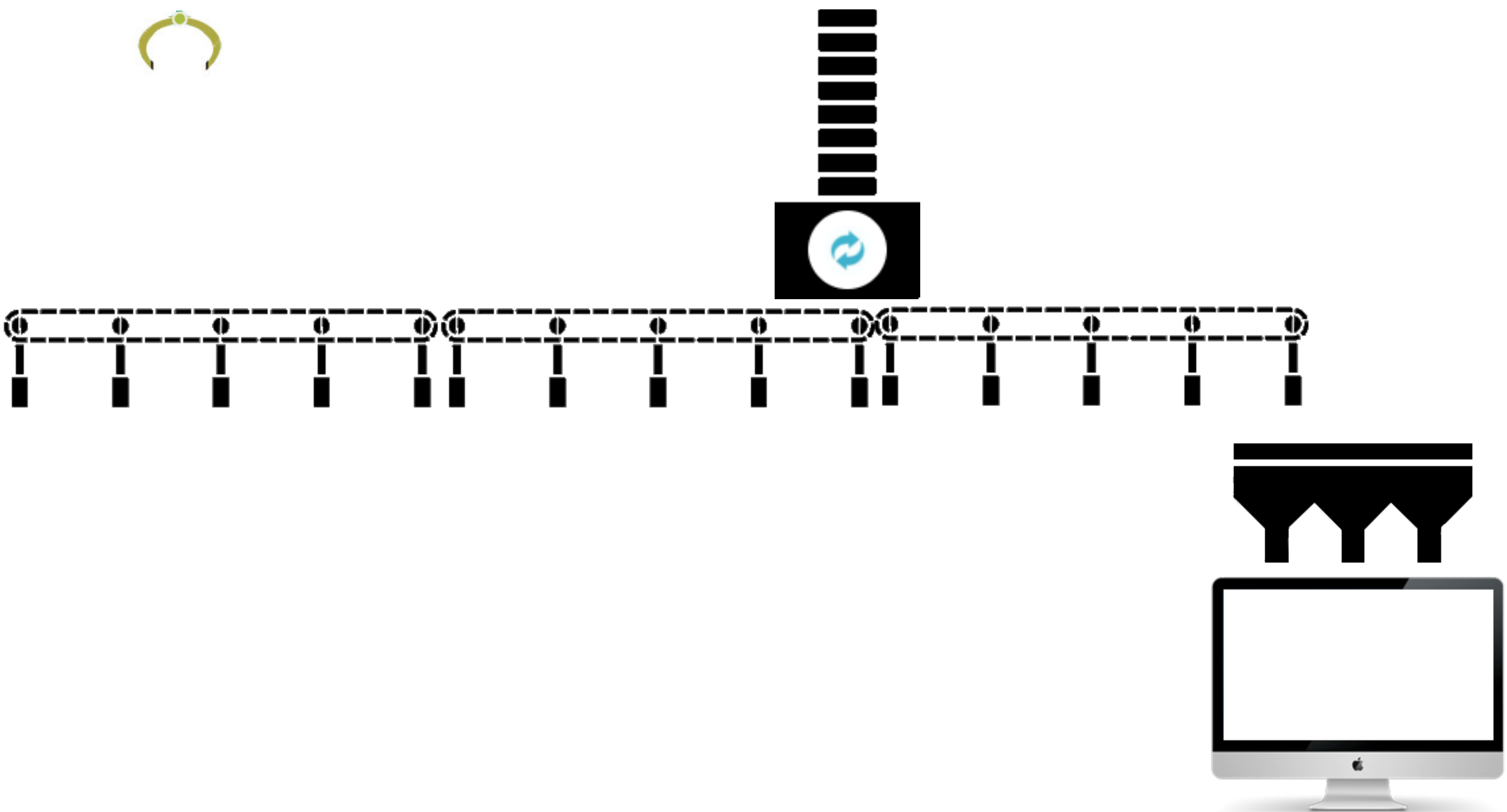


SUPERVISED LEARNING

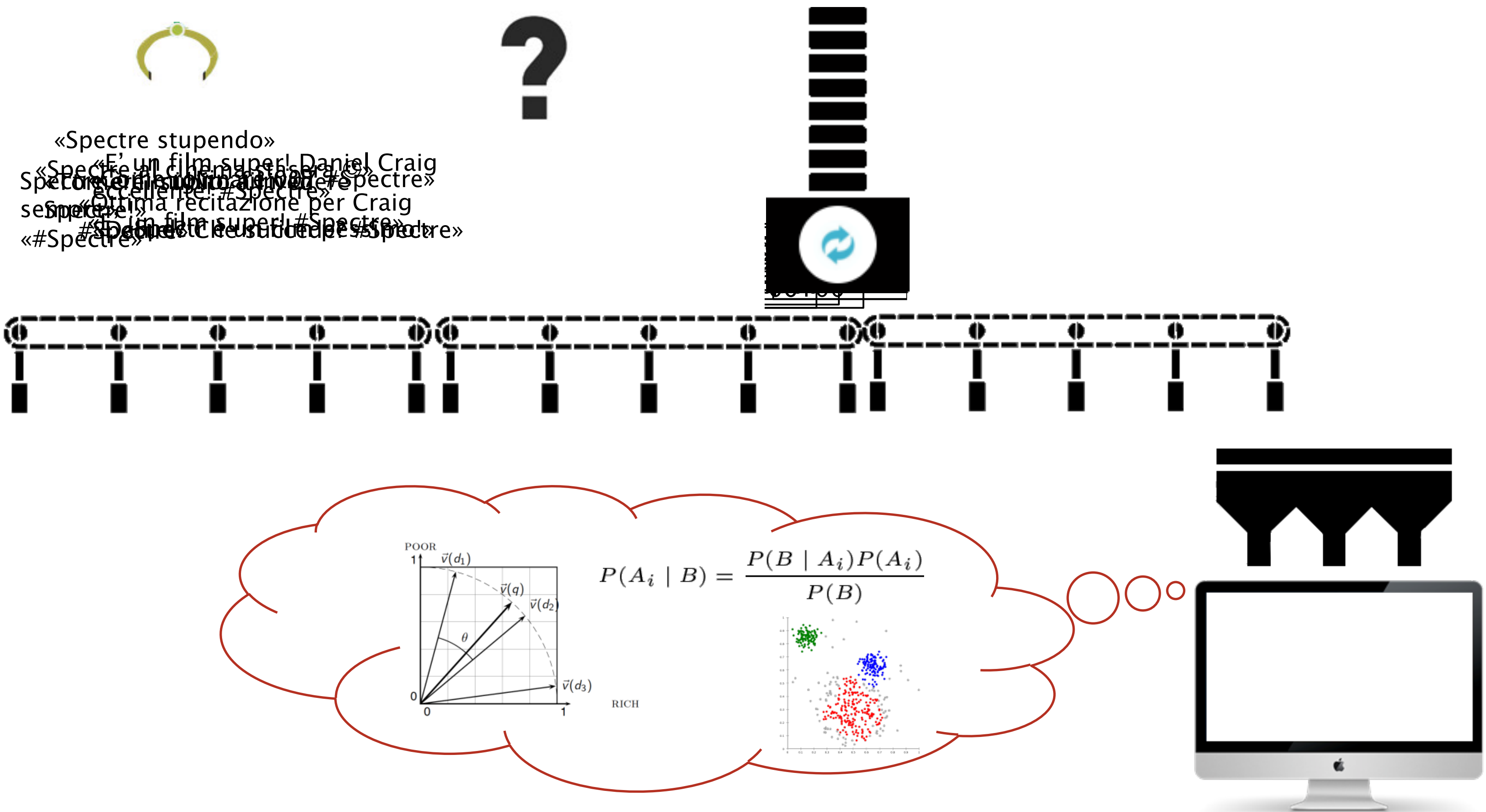
«Spectre si è rivelato un film deludente!»



SUPERVISED LEARNING



SUPERVISED LEARNING



SUPERVISED LEARNING

- How to handle **negation**
  - `I` **`didn't`** `like this movie`
    vs
  - `I really like this movie`
- Which words to use?
  - Only adjectives
  - All words
    - All words turns out to work better, at least on this data

SUPERVISED LEARNING

Add NOT_ to every word between **negation** and following punctuation:

```
didn’t like this movie , but I
```

```
didn’t NOT_like NOT_this NOT_movie but I
```

Das, Sanjiv and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific Finance Association Annual Conference.
Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.

SEMI-SUPERVISED LEARNING (1)

- Use a small amount of information
  - A few labeled examples
  - A few hand-built patterns
  - Distant supervision (emotico, emoji, tags, etc..)

- To bootstrap a lexicon

SEMI-SUPERVISED LEARNING (1)

- Adjectives conjoined by "*and*" have same polarity
  - Fair **and** legitimate
  - corrupt **and** brutal
- Adjectives conjoined by "***but***" do not
  - fair **but** brutal

Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the Semantic Orientation of Adjectives. ACL, 174–181

SEMI-SUPERVISED LEARNING (1)

- Label **seed set** of 1336 adjectives (all >20 in 21 million word WSJ corpus)
  - 657 positive
    - *adequate central clever famous intelligent remarkable reputed sensitive slender thriving…*
  - 679 negative
    - *contagious drunken ignorant lanky listless primitive strident troublesome unresolved unsuspecting…*

SEMI-SUPERVISED LEARNING (1)

- Expand seed set to conjoined adjectives

Google "was nice and"

Nice location in Porto and the front desk staff **was nice and** helpful ...
www.tripadvisor.com/ShowUserReviews-g189180-d206904-r12068... +1
Mercure Porto Centro: Nice location in Porto and the front desk staff **was nice and** helpful - See traveler reviews, 77 candid photos, and great deals for Porto, ...

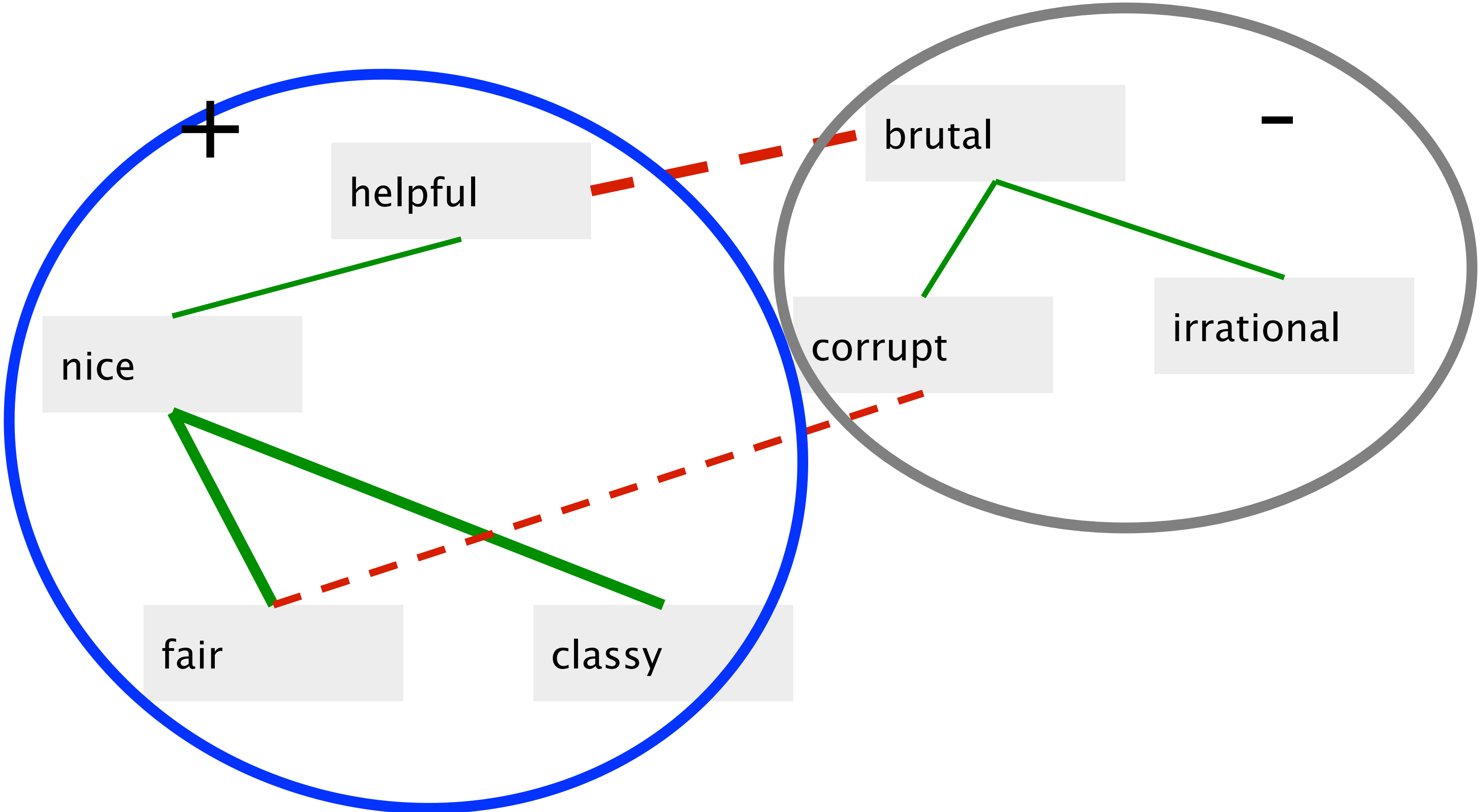
nice, helpful

If a girl **was nice and** classy, but had some vibrant purple dye in ...
answers.yahoo.com › Home › All Categories › Beauty & Style › Hair +1
4 answers - Sep 21
Question: Your personal opinion or what you think other people's opinions might ...
Top answer: I think she would be cool and confident like katy perry :)

nice, classy

SEMI-SUPERVISED LEARNING (1)

- Clustering for partitioning the graph into two



SEMI-SUPERVISED LEARNING (1)

- Positive
  - *bold decisive disturbing generous good honest important large mature patient peaceful positive proud sound stimulating straightforward strange talented vigorous witty…*
- Negative
  - *ambiguous cautious cynical evasive harmful hypocritical inefficient insecure irrational irresponsible minor outspoken pleasant reckless risky selfish tedious unsupported vulnerable wasteful…*

SEMI-SUPERVISED LEARNING (2)

1. Extract a ***phrasal lexicon*** from reviews
2. **Learn polarity** of each phrase
3. Rate a review by the **average polarity** of its phrases

Turney (2002): Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews

SEMI-SUPERVISED LEARNING (2)

First Word	Second Word	Third Word (not extracted)
JJ	NN or NNS	anything
RB, RBR, RBS	JJ	Not NN nor NNS
JJ	JJ	Not NN or NNS
NN or NNS	JJ	Nor NN nor NNS
RB, RBR, or RBS	VB, VBD, VBN, VBG	anything

SEMI-SUPERVISED LEARNING (2)

- How to measure polarity?
  - **Positive** phrases co-occur more with "*excellent*"
  - **Negative** phrases co-occur more with "*poor*"
  - But **how to measure co-occurrence**?

SEMI-SUPERVISED LEARNING (2)

- **Mutual information** between 2 random variables X and Y

$$\mathrm{I}(X,Y) = \sum_{x}\sum_{y} P(x,y)\log_2 \frac{P(x,y)}{P(x)P(y)}$$

- **Pointwise mutual information**:
  - How much more do events x and y co-occur than if they were independent?

$$\mathrm{PMI}(X,Y) = \log_2 \frac{P(x,y)}{P(x)P(y)}$$

SEMI-SUPERVISED LEARNING (2)

- **PMI between two words**:
  - How much more do two words co-occur than if they were independent?

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

SEMI-SUPERVISED LEARNING (2)

- Query search engine
  - P(word) estimated by `hits(word)/N`
  - P($word_1$,$word_2$) by `hits(word1 NEAR word2)/N`

$$\mathrm{PMI}(word_1, word_2) = \log_2 \frac{\frac{1}{N} hits(word_1 \text{ NEAR } word_2)}{\frac{1}{N} hits(word_1) \frac{1}{N} hits(word_2)}$$

SEMI-SUPERVISED LEARNING (2)

- Does phrase appear more with "**poor**" or "**excellent**"?

$$\text{Polarity}(\textit{phrase}) = \text{PMI}(\textit{phrase}, \text{"excellent"}) - \text{PMI}(\textit{phrase}, \text{"poor"})$$

$$= \log_2 \frac{\frac{1}{N}\text{hits}(\textit{phrase}\ \text{NEAR}\ \text{"excellent"})}{\frac{1}{N}\text{hits}(\textit{phrase})\frac{1}{N}\text{hits}(\text{"excellent"})} - \log_2 \frac{\frac{1}{N}\text{hits}(\textit{phrase}\ \text{NEAR}\ \text{"poor"})}{\frac{1}{N}\text{hits}(\textit{phrase})\frac{1}{N}\text{hits}(\text{"poor"})}$$

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

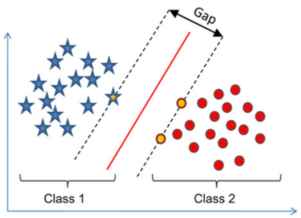
SEMI-SUPERVISED LEARNING (2)

Phrase	POS tags	Polarity
direct deposits	JJ NNS	5.8
online web	JJ NN	1.9
very handy	RB JJ	1.4
…		
virtual monopoly	JJ NN	-2.0
lesser evil	RBR JJ	-2.3
other problems	JJ NNS	-2.8
low funds	JJ NNS	-6.8
unethical practices	JJ NNS	-8.5
Average		**-1.2**

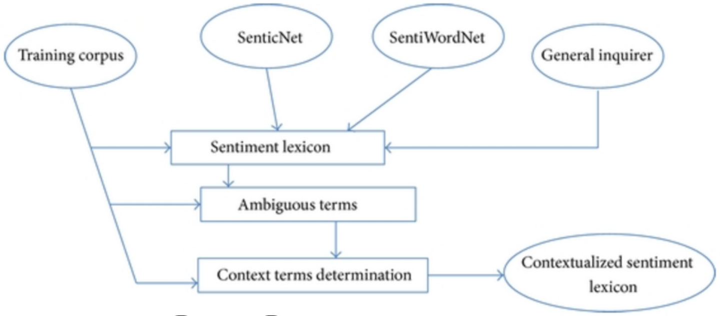
CLASSIFY

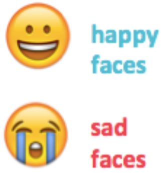
- Lexicon-based approach

- Supervised learning

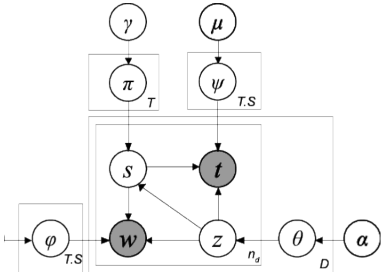


- Semi-Supervised learnin





- Unsupervised learning



UNSUPERVISED LEARNING

- Unsupervised learning is suitable for **aspect-based** sentiment analysis

"Ho comprato il nuovo iPhone. Ha un bellissimo design caratterizzato da colori molto vivi. Unica pecca è la pessima batteria"

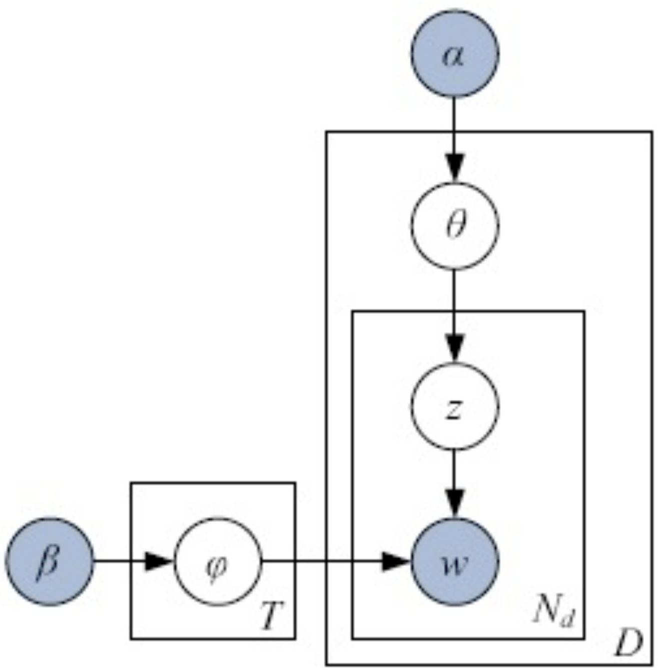
ASPECT: design
SENTIMENT: positivo (bellissimo)

ASPECT: batteria
SENTIMENT: negativo (pessima)

UNSUPERVISED LEARNING

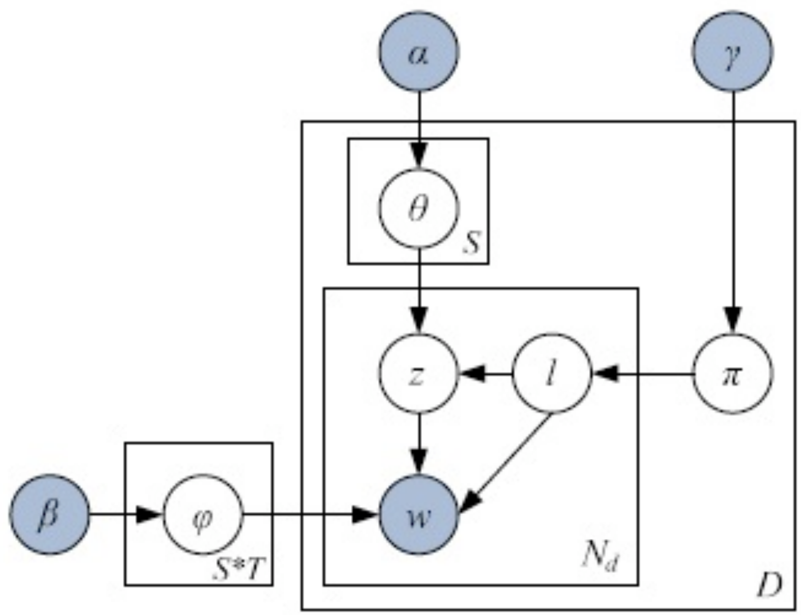
- **Joint Sentiment Topic** model
  - Model base on <u>Latent Dirichlet Allocation</u>

LDA



(a)

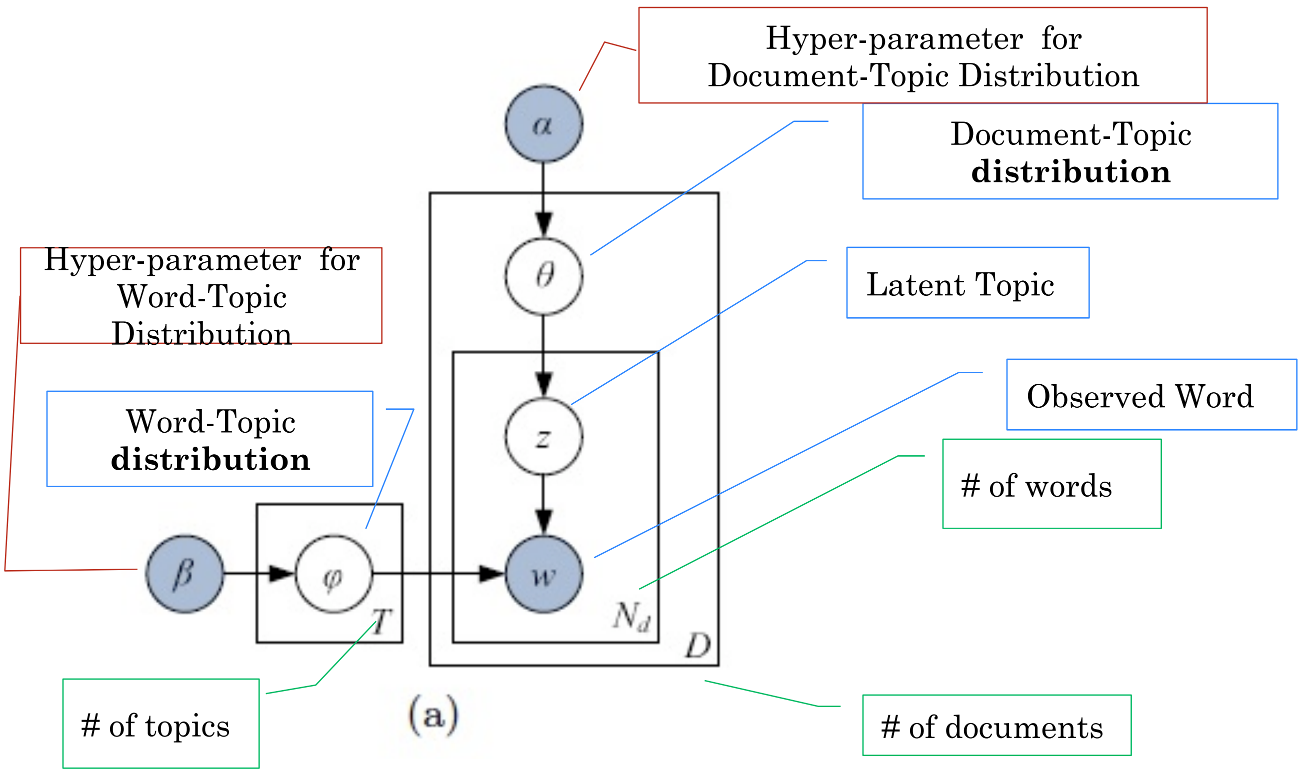
JST



(b)

UNSUPERVISED LEARNING

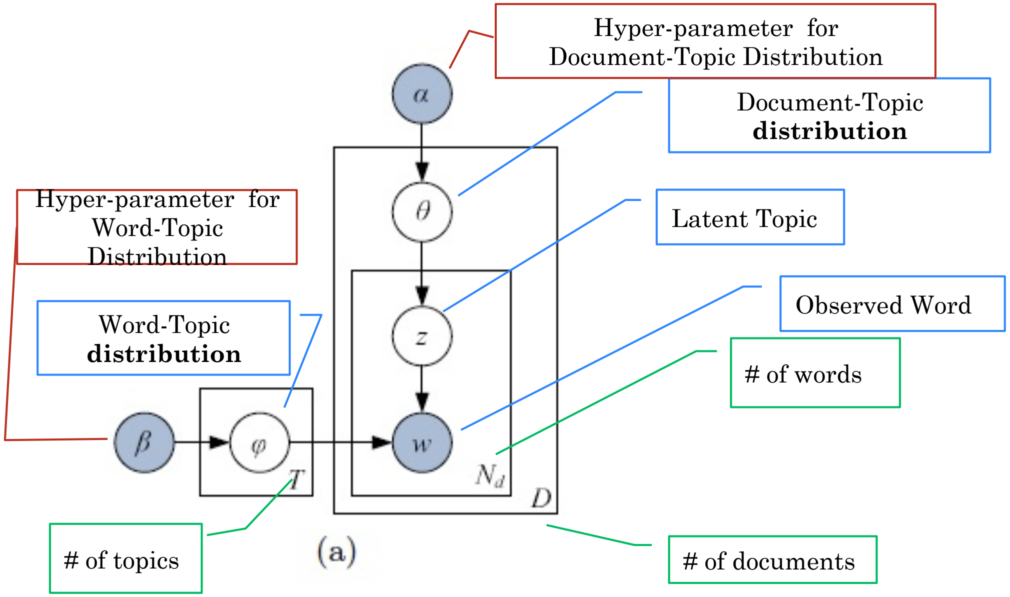
- **<u>Latent Dirichlet Allocation</u>**



UNSUPERVISED LEARNING

Given the parameters α and β, the joint distribution of a topic mixture θ, a set of topics z, and a set of N words w is given by:

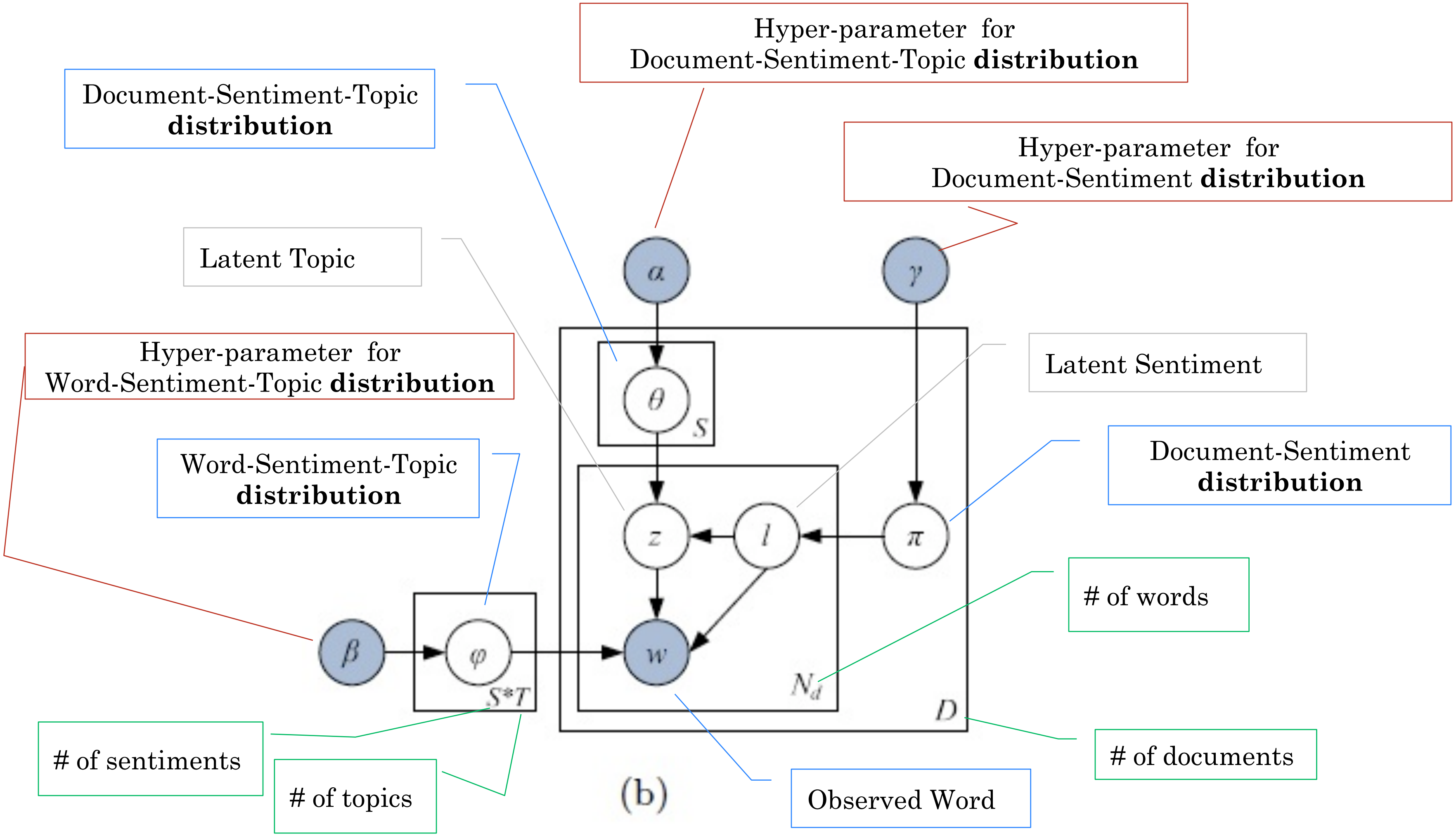
$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta)$$



(a)

UNSUPERVISED LEARNING

- **Joint Topic/Sentiment (JST)**

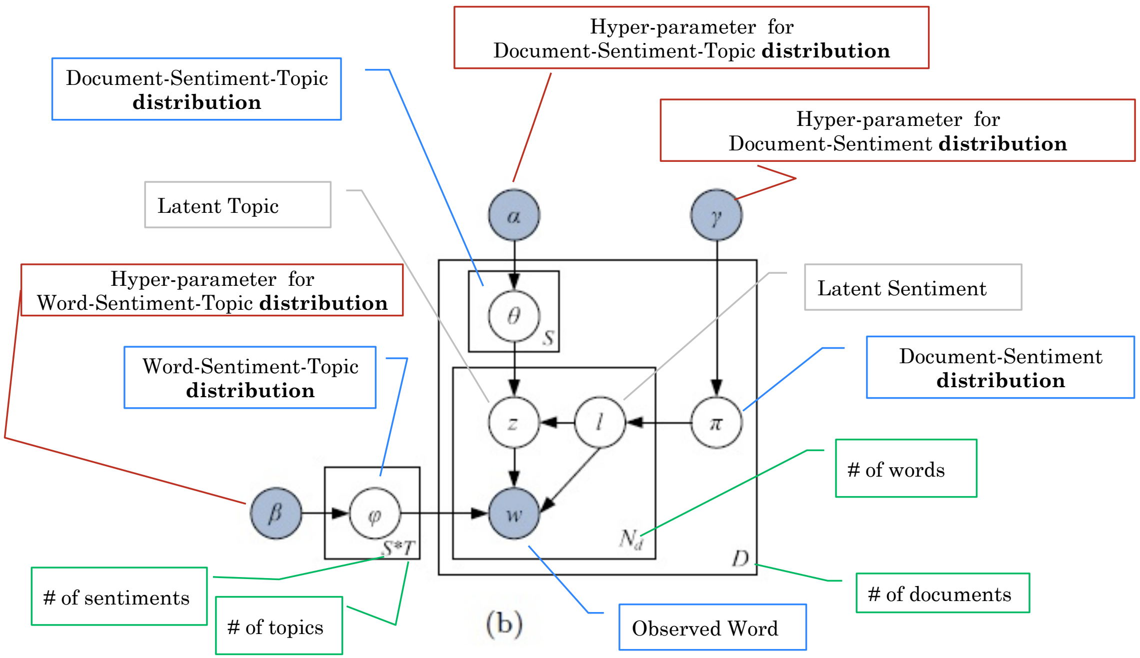


(b)

UNSUPERVISED LEARNING

- **Joint Topic/Sentiment (JST)**
  - The joint probability of the topic/sentiment label assignments and the words can be factored into the following three terms:

$$P(\mathbf{w}, \mathbf{z}, \mathbf{l}) = P(\mathbf{w}|\mathbf{z}, \mathbf{l})P(\mathbf{z}|\mathbf{l}, d)P(\mathbf{l}|d)$$



UNSUPERVISED LEARNING

- **Joint Topic/Sentiment (JST)**
  - Gibbs sampling will sequentially sample each variable of interest, z_t and l_t, from the distribution over that variable given the current values of all other variables and the data. The conditional posterior for z_t and l_t is:

$$P(z_t = j, l_t = k | \mathbf{w}, \mathbf{z}_{-t}, \mathbf{l}_{-t}, \alpha, \beta, \gamma) \propto$$

$$\boxed{\frac{\{N_{w_t,j,k}\}_{-t}+\beta}{\{N_{j,k}\}_{-t}+V\beta}} \cdot \boxed{\frac{\{N_{j,k,d}\}_{-t}+\alpha}{\{N_{k,d}\}_{-t}+T\alpha}} \cdot \boxed{\frac{\{N_{k,d}\}_{-t}+\gamma}{\{N_d\}_{-t}+S\gamma}}$$

$\boxed{P(\mathbf{w}|\mathbf{z},\mathbf{l})}$ $\boxed{P(\mathbf{z}|\mathbf{l},d).}$ $\boxed{P(\mathbf{l}|d)}$

GIBBS SAMPLING

UNSUPERVISED LEARNING

- **Joint Topic/Sentiment (JST)**
  - We can make use of some prior knowledge, e.g. **lexicons** of **positive**, **negative** and **neutral** orientations

$$P(z_t = j, l_t = k \mid \mathbf{w}, \mathbf{z}_{-t}, \mathbf{l}_{-t}, \alpha, \beta, \gamma) \propto$$

$$\boxed{\frac{\{N_{w_t,j,k}\}_{-t} + \beta}{\{N_{j,k}\}_{-t} + V\beta}} \cdot \boxed{\frac{\{N_{j,k,d}\}_{-t} + \alpha}{\{N_{k,d}\}_{-t} + T\alpha}} \cdot \boxed{\frac{\{N_{k,d}\}_{-t} + \gamma}{\{N_{d}\}_{-t} + S\gamma}}$$

$$\boxed{P(\mathbf{w} \mid \mathbf{z}, \mathbf{l})} \qquad \boxed{P(\mathbf{z} \mid \mathbf{l}, d)} \qquad \boxed{P(\mathbf{l} \mid d)}$$

compare each word w_t against the lexicons..**if** there is a match, w_i is assigned to the corresponding **sentiment** label

otherwise

GIBBS SAMPLING

UNSUPERVISED LEARNING

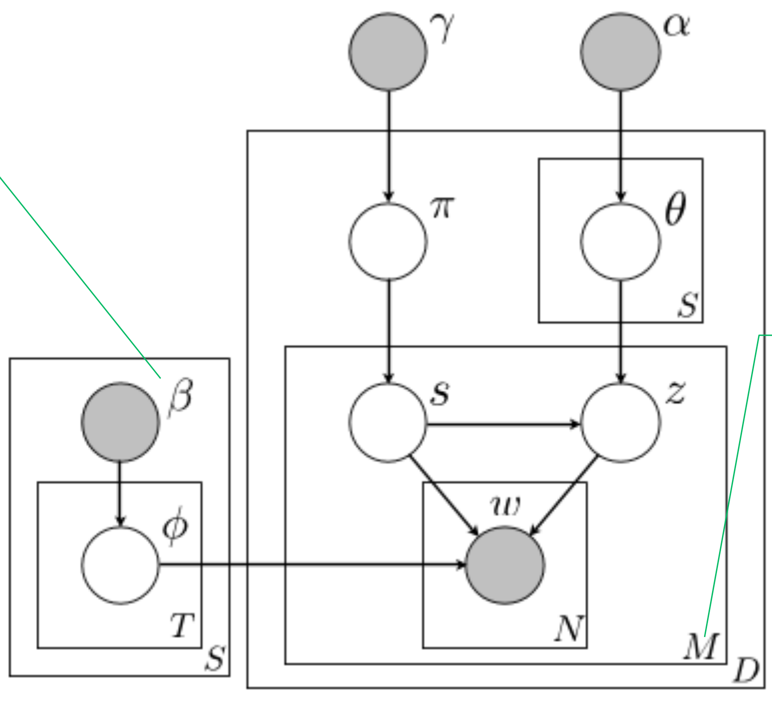
- **Joint Topic/Sentiment (JST)**
  - What does it produce?

Positive sentiment label						Negative sentiment label					
Topic 1		Topic 2		Topic 3		Topic 1		Topic 2		Topic 3	
w	$P(w\|z,l)$	w	$P(w\|z,l)$	w	$P(w\|z,l)$	w	$P(w\|z,l)$	w	$P(w\|z,l)$	w	$P(w\|z,l)$
good	0.084708	tom	0.035175	ship	0.059020	bad	0.079132	sex	0.065904	prison	0.073208
realli	0.046559	ryan	0.030281	titan	0.031586	worst	0.035402	scene	0.053660	evil	0.032196
plai	0.044174	hank	0.025388	crew	0.024439	plot	0.033687	sexual	0.031693	guard	0.031755
great	0.036645	comedi	0.021718	cameron	0.024439	stupid	0.029767	women	0.026291	green	0.029109
just	0.028990	star	0.020800	alien	0.022826	act	0.025602	rate	0.023770	hank	0.028227
perform	0.028362	drama	0.016519	jack	0.020751	suppos	0.025480	act	0.023230	wonder	0.027345
nice	0.026354	meg	0.015601	water	0.019137	script	0.024500	offens	0.018728	excute	0.026904
fun	0.025978	joe	0.014378	stori	0.017984	wast	0.024500	credict	0.016027	secret	0.025581
lot	0.025853	relationship	0.014072	rise	0.016601	dialogu	0.023643	porn	0.014587	mile	0.022936
act	0.022715	mail	0.013766	rose	0.013835	bore	0.022908	rape	0.013867	death	0.022495
direct	0.021586	blond	0.013460	boat	0.013374	poor	0.022908	femal	0.013686	base	0.022054
best	0.020331	run	0.012543	deep	0.013143	complet	0.020825	cut	0.013686	tom	0.019849
get	0.020331	phone	0.012237	ocean	0.012451	line	0.019968	gril	0.013506	convict	0.018967
entertain	0.018198	date	0.011931	board	0.011990	terribl	0.018988	parti	0.012426	return	0.018526
better	0.017445	got	0.011625	sink	0.011299	mess	0.015313	male	0.011886	franklin	0.016762
job	0.016692	busi	0.011319	sea	0.010838	wors	0.014338	bad	0.011346	happen	0.016321
talent	0.016064	cute	0.011013	rain	0.010838	dull	0.013598	nuditi	0.011166	power	0.014116
pretti	0.016064	sister	0.010708	dicaprio	0.010607	actor	0.012986	woman	0.010986	known	0.012352
try	0.015688	children	0.010096	storm	0.010377	total	0.012986	peopl	0.010986	instinct	0.011470
want	0.015186	dog	0.009790	disast	0.010146	isn	0.012863	nake	0.010625	inmat	0.011470

UNSUPERVISED LEARNING

- **Aspect Sentiment Unification Model (ASUM)**
  - Similar to JST, but…

β is different for positive and negative!



It assumes that a document is composed of **M sentences**..one sentence describes one aspect!

SEMANTIC SENTIMENT ANALYSIS

- **Semantic sentiment analysis** aims at extracting and using the underlying **semantics of words** in identifying their sentiment orientation with regards to their **context** in the text.
  - It relies on
    - **external semantic knowledge bases** (e.g. ontologies and semantic networks)

AND

    - **NLP techniques** to capture the conceptual representations of words that implicitly convey sentiment

SEMANTIC SENTIMENT ANALYSIS

- ***Sentic Computing***
  - It employs **two knowledge-base sources** of common sense concepts and affective information
    - **ConceptNet:** a semantic graph representation of the common sense information in the Open Mind corpus. Nodes in the graph represents concepts and edges represents the assertions of common-sense that interconnect concepts.
    - **WordNet-Affect:** a linguistic resource of affective knowledge built upon WordNet

SEMANTIC SENTIMENT ANALYSIS

- ***Sentic Computing***
  - ConceptNet and WordNet-Affect are used to extract
    - the common-sense concepts (e.g., “nice day”, “simple life”)
    - with their associated semantics and sentics (Hourglass of Emotions)
    - break text into clauses and, hence, deconstruct such clauses into small bags of concepts, in order to feed these into a commonsense reasoner

SEMANTIC SENTIMENT ANALYSIS

- ***Sentic Computing***

1. **Breaks text into clauses**: each verb and its associated noun phrase are considered in turn

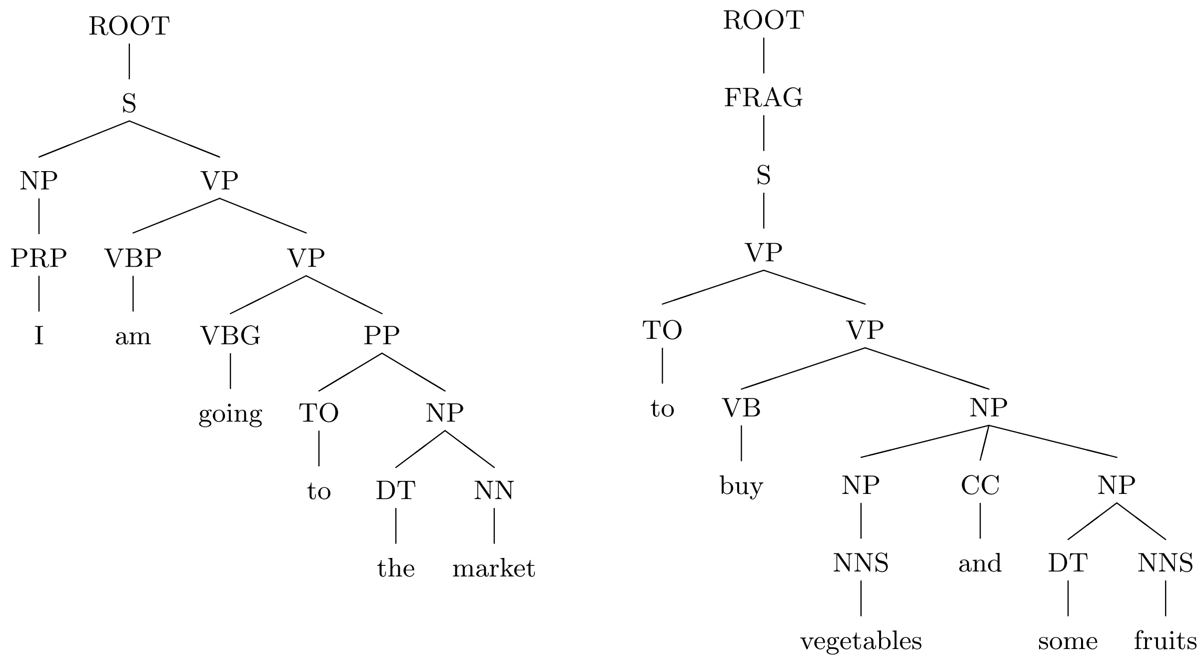
"I am going to the market to buy vegetables and some fruits"

"I am going to the market" + *"market to buy vegetables and some fruits"*

SEMANTIC SENTIMENT ANALYSIS

- *Sentic Computing*

2. **Separates clauses** into verb and noun chunks



SEMANTIC SENTIMENT ANALYSIS

- *Sentic Computing*

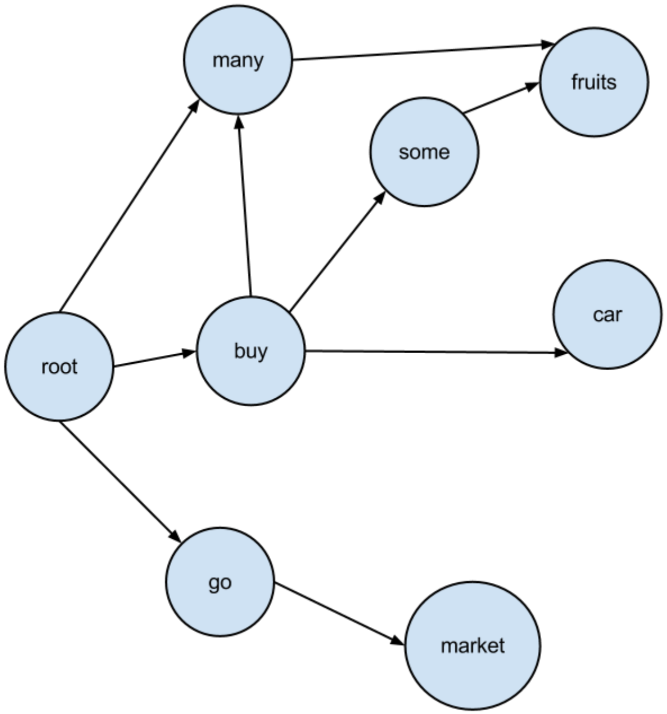
3. **Create POS pairs:**
    - ADJECTIVE-NOUN: both taken as concepts
    - NOUN-ADJECTIVE: adjective discarded , noun taken as concept
    - NOUN-NOUN : both taken single concept
    - NOUN-STOPWORD (and viceversa) : stopword discarded, noun taken

“market, some fruits, fruits, and vegetables”

SEMANTIC SENTIMENT ANALYSIS

- *Sentic Computing*

4. **Create parse graph and search for "event concepts":** In order to capture event concepts, matches between the concepts and the stemmed verb chunks are searched.



go market
buy some fruits
buy fruits
and buy vegetables
...

SEMANTIC SENTIMENT ANALYSIS

- ***Sentic Computing***

5. Aligning ConceptNet with WordNet-Affect:

- Take the WordNet-Affect matrix (14.301 concepts × 117.365)

AffectNet	*IsA-pet*	*KindOf-food*	*Arises-joy*	...
dog	0.981	0	0.789	...
cupcake	0	0.922	0.910	...
songbird	0.672	0	0.862	...
gift	0	0	0.899	...
sandwich	0	0.853	0.768	...
rotten fish	0	0.459	0	...
win lottery	0	0	0.991	...
bunny	0.611	0.892	0.594	...
police man	0	0	0	...
cat	0.913	0	0.699	...
rattlesnake	0.432	0.235	0	...
...	...	...	...	...

- Align lemma of event concepts with lemma of the words in WordNet-Affect. The degree of similarity between two concepts is the dot product between their rows in A.
- Apply a singular value decomposition (SVD) on the resulting matrix and use dimensionality reduction to discard those components representing relatively small variations in the data.

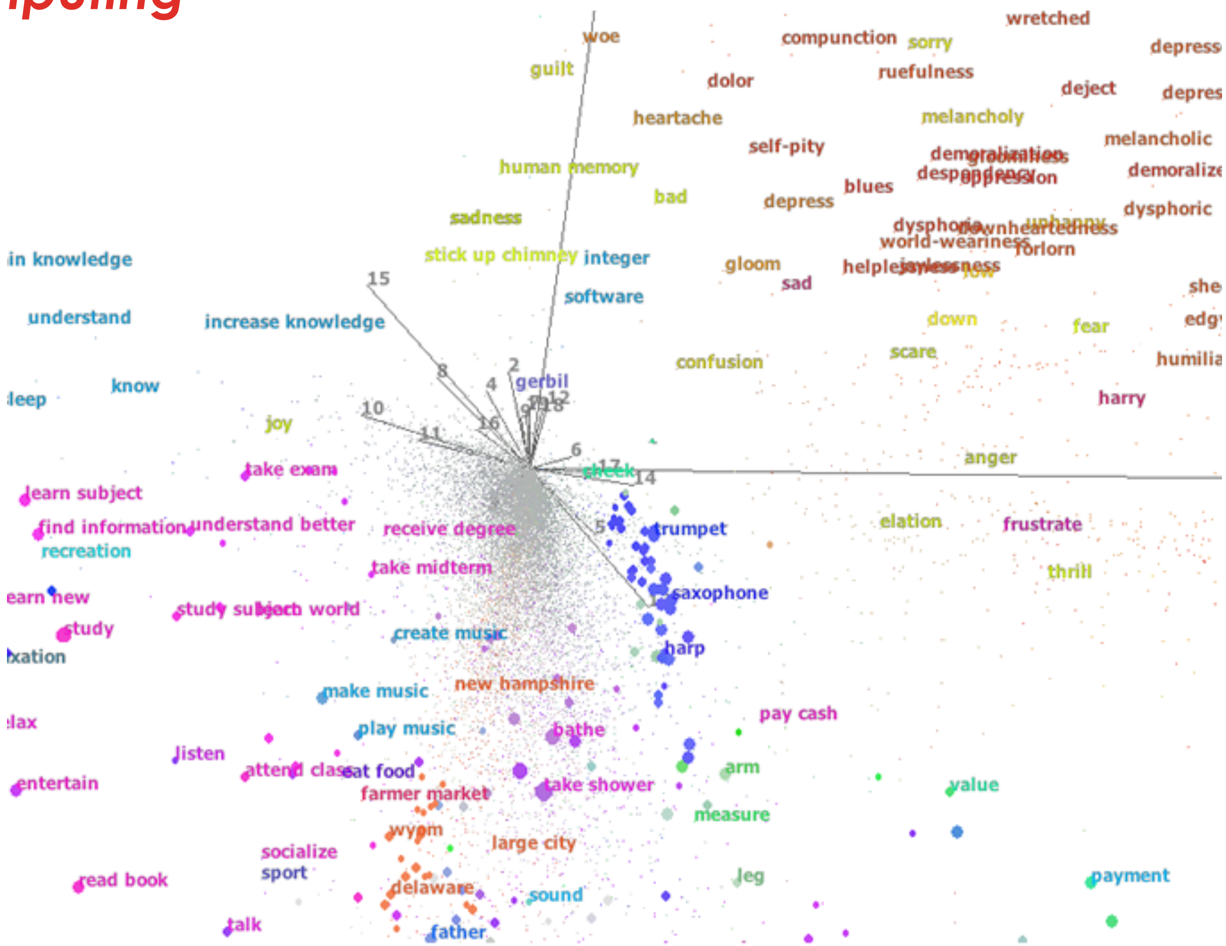
SEMANTIC SENTIMENT ANALYSIS

- ***Sentic Computing***

5. **Aligning ConceptNet with WordNet-Affect:**
    - If we choose to discard all but the first *k* principal components, commonsense concepts are represented by vectors of *k* coordinates, which can be seen as describing multi-word expressions in terms of 'eigenconcepts'

SEMANTIC SENTIMENT ANALYSIS

- *Sentic Computing*



SEMANTIC SENTIMENT ANALYSIS

- *Sentic Computing*

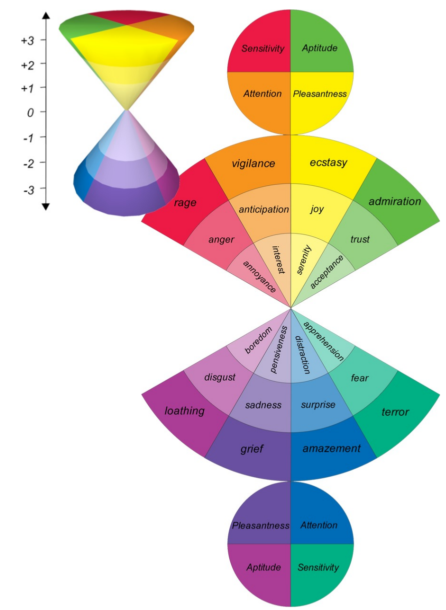
5. **Aligning ConceptNet with WordNet-Affect:**
   - If we choose to discard all but the first *k* principal components, commonsense concepts are represented by vectors of *k* coordinates, which can be seen as describing multi-word expressions in terms of 'eigenconcepts'
   - In order to measure such semantic relatedness, then, A'=SVD(A) is clustered by using a k-medoid approach.
   - Once a new concept is extracted, it is assigned to the most similar medoid.

PROBLEM: SVD reduction technique on a very large and sparse matrix can have high computational complexity

SEMANTIC SENTIMENT ANALYSIS

- **<u>SOLUTION:</u>** ***SenticNet***
  - Generate the AffectiveSpace for once and store its common sense concepts along with their sentics in a static lexicon*
  - **<u>Concept polarity</u>** is the algebraic sum of the Hourglass model's sentic labels.

$$p_c = \frac{Plsn(c) + |Attn(c)| - |Snst(c)| + Aptt(c)}{9}$$



*This allows to extract the sentic information of a given common sense concept without the need for regenerating the AffectiveSpace

DIFFICULTIES

- Ambiguity

TENERE

"Oh, queste cagnoline sono proprio **tenere**!"

"Devo **tenere** a mente i miei prossimi impegni."

- Irony and sarcasm

FIGURES OF SPEECH

- IRONY:
  - Irony is a figure of speech in which words are used in such a way that their intended meaning is the opposite from the actual meaning of the words especially in order to **be funny**

"Oh great! Now you have broken my new camera."

- SARCASM:
  - Sarcasm is a literary and rhetorical device that is meant to mock with often satirical remarks with a typical purpose to **hurt someone**

"Some cause happiness wherever they go; others whenever they go."

STATE OF THE ART

- Mostly focused on training baseline classifiers with several linguistic characteristics to distinguish irony vs not-irony and sarcasm vs not-sarcasm

	Linguistic Features	**Learning Scheme**
[Davidov et al. 2010]	unigrams, punctuation	KNN
[Reyes et al. 2013]	emoticons, quotes, capitalized words, adverbs, divergences related to verbs, semantic similarity, n-grams, skip-grams, polarity s-grams, activation, imagery, and pleasantness	NB, DT
[Ptácek et al. 2014]	n-grams, patterns, POS tags, emoticon, punctuation, word case	MaxEnt, SVM
[Barbieri and Saggion, 2014]	Frequency, Written-Spoken, Structure, Intensity, Synonyms, Ambiguity, Sentiments	DT, RF
[Fersini et al., 2015]	emoticons, emphatic expressions, onomatopoeic expressions, punctuation, POS tags	Ensemble Models
[Hernandez-Farias et al., 2015, 2016]	statistical-based and lexicon-based features	NB, DT, MAxEnt, SVM, MLP

STATE OF THE ART

- Recently some effort has been spent on
  - **machine learning** perspective: combining several learners

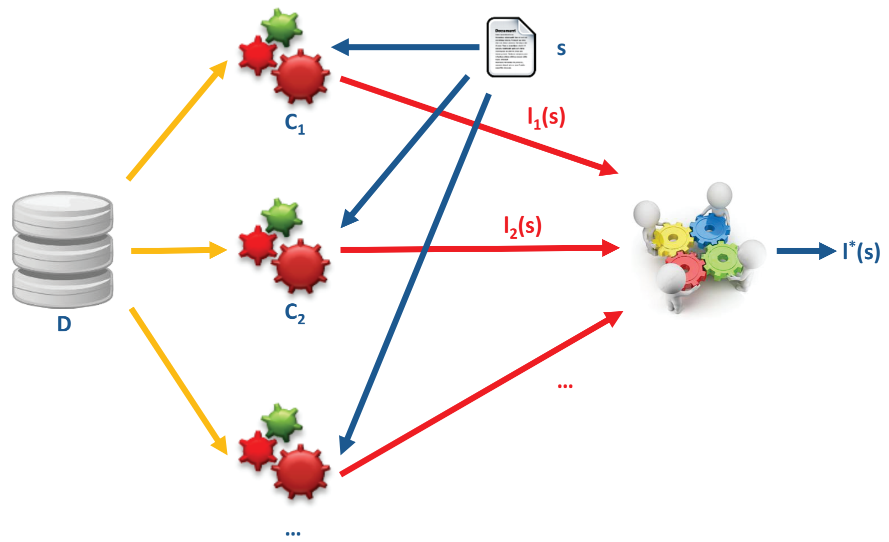
    Ensemble methods?

  - understanding **which linguistic feature contributes more** on the characterization of sarcasm and irony

    Are all features equally reliable?

LET'S START FROM THE ENSEMBLE

- The most pupular schema:
  - Majority Voting (MV)



- **Limitation**:
  - the models have **uniform** distributed **weights** regardless their reliability.
    - when dealing with sarcasm and irony, an ensemble should take into account the reliability of each model to detect **rare statements**

BAYESIAN MODEL AVERAGING

- Given a sentence **s** and a set **C** of independent classifiers producing their prediction l(s), the optimal label *l*(s)* is determined as:

$$l^*(s) = \arg\max_{l(s)} P(l(s) \mid C, \mathcal{D}) \quad = \quad \sum_{i \in C} \underbrace{P(l(s) \mid i, \mathcal{D})}_{\text{Marginal distribution}} \underbrace{\boxed{P(\mathcal{D} \mid i)}}_{\text{Model reliability}}$$

$$\boxed{P(\mathcal{D} \mid i) \approx \frac{1}{\iota} \sum_{\iota=1}^{\phi} \frac{2 \times P_{i\iota}(\mathcal{D}) \times R_{i\iota}(\mathcal{D})}{P_{i\iota}(\mathcal{D}) + R_{i\iota}(\mathcal{D})}}$$

BAYESIAN MODEL AVERAGING

Optimal ensemble composition

- Including many powerful classifiers in an ensemble does not ensure a good ensemble!
- A good ensemble (S) should minimize three main types of error :
  - **Bias**: error caused by an inaccurate model

$$E[S] = \frac{1}{|D|}\sum_{m=1}^{|D|} f_e(l(m)^*, l(m)^{BMA})]$$

  - **Variance**: error caused by the data sample

$$Var(S) = \frac{1}{n^2} Var\Big(\sum_{i\in S} i\Big) = $$
$$= \frac{1}{n^2}\Big[\Big(\sum_{i\in S} Var(i)\Big) + 2\sum_{i\in S}\sum_{\substack{j<i\\ j\in S}} Cov(i,j)\Big]$$

  - **Noise**: (unpredictable) error on unseen samples

BAYESIAN MODEL AVERAGING

Optimal ensemble composition

- What counts for determining the optimal ensemble is the **trade-off** given by:

$$Trade-Off = E[S] + Var(S)$$

- **Bias**: the higher is the prediction accuracy of the ensemble and the lower is the bias.
- **Variance**: decreases as the number of independent classifiers increases
  - If classifiers are positively correlated, the variance increase for increasing number of models.

Reduce bias and variance trough an **optimal ensemble composition**!

BAYESIAN MODEL AVERAGING

Optimal ensemble composition

- Combinatorial optimization problem: $\sum_{p=1}^{n} \frac{n!}{p!(n-p)!}$ possible compositions!
- Proposal:
  - compute the **Discriminative Marginal Contribution (DMC)** of each classifier

    a. gain that classifier i gives with respect to classifier j:

    ① j incorrectly labels the sentence s, but i correctly tags it

    ② both i and j correctly label s

    b. loss that classifier introduces with respect to classifier j:

    ③ j correctly labels sentence s, but i incorrectly tags it

    ④ both i and j incorrectly label s

BAYESIAN MODEL AVERAGING

Optimal ensemble composition

- The contribution r_i^S of each classifier *i* belonging to a given ensemble $S \subseteq C$ can be estimated as:

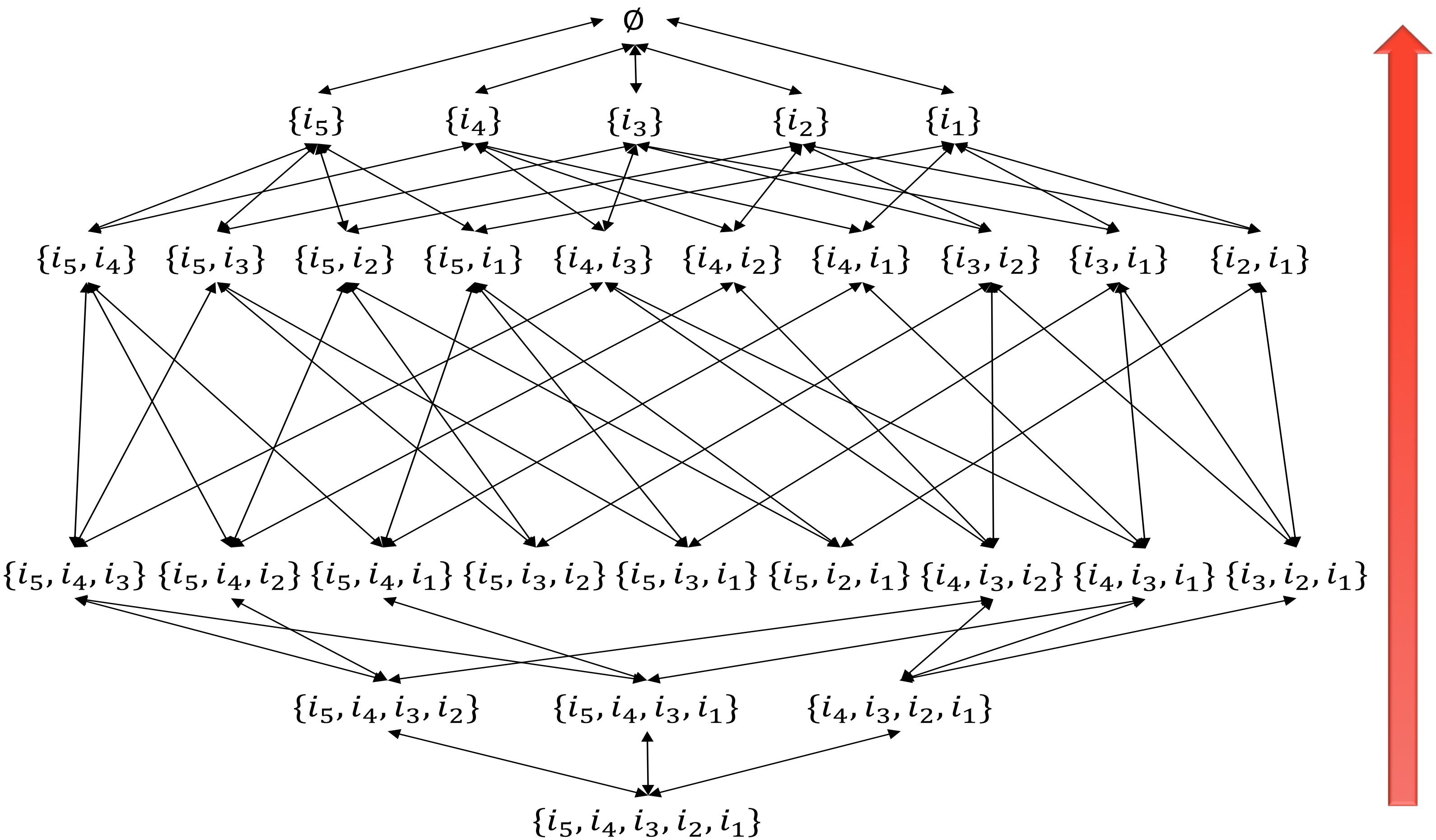
gain

$$r_i^S = \frac{\sum_{j \in \{S \setminus i\}} \sum_{q \in \{0,1\}} P(i = 1 \mid j = q) P(j = q)}{\sum_{j \in \{S \setminus i\}} \sum_{q \in \{0,1\}} P(i = 0 \mid j = q) P(j = q)}$$

loss

- **Greedy strategy**: **backward selection** until a local maxima of average classifier contribution

BAYESIAN MODEL AVERAGING



LINGUISTIC FEATURES

- Which feature is **more discriminative** for irony? And for sarcasm?
  - Bag of words **BOW**
  - Pragmatic Particles **PP**
    - Emoticons ☺
    - Initialisms for emphatic expressions *ROFL*
    - Onomatopoeic expressions *bleh*
    - Punctuation ????!!!!.....
  - Part-Of-Speech **POS**
  - Combination **PP&POS**

Nominal	**Other closed-class words**
N - common noun O - pronoun (personal; not possessive) ^ - proper noun S - nominal + possessive Z - proper noun + possessive	D - determiner P - pre- or postposition, or subordinating conjunction & - coordinating conjunction T - verb particle X - existential there
Other open-class words	**Other compounds**
V - verb A - adjective R - adverb ! - interjection	L - nominal + verbal (e. g., i'm) M - proper noun + verbal Y - X + verbal

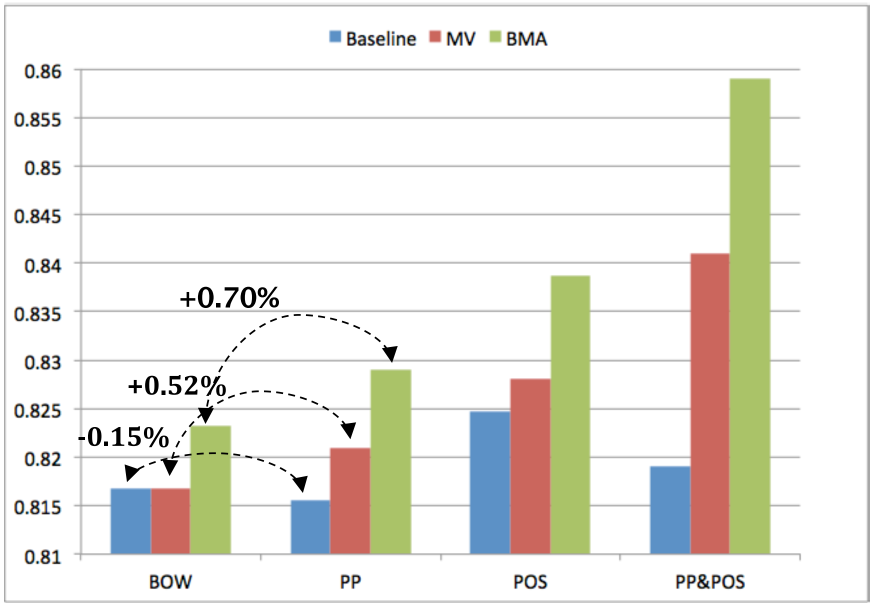
EXPERIMENTAL SETTINGS

- Sarcasm: originally 12.345 tweets
  - Reduced to 8,000 according to the "2over3" annotator agreement*
  - Randomly sampled in 2 datasets composed of 4,000 tweets each containing
- Irony: 40,000 tweets [Reyes at al. 2013]
  - 10,000 ironic vs 30,000 non-ironic tweets (distinguished in Education, Politics, and Humor)
- Baseline Classifiers (also included in BMA):

Multinomial Naïve Bayes (MNB)	Bayesian Networks (BN)	Majority Voting (MV)
Support Vector Machines (SVM)	Decision Trees (DT)	

- Performance measures: + for ironic/sarcastic, - for non-ironic/non-sarcastic
  - Accuracy, Precision(+), Recall(+), F-Measure(+), Precision(-), Recall(-), F-Measure(-), Bias-Variance Trade-Off

RESULTS ON **SARCASM**: LINGUISTIC FEATURES



BMA	BOW	PP	POS	PP & POS
P+	**0.7709**	0.7718	0.7605	0.7697
R+	0.8245	0.8375	0.8540	**0.8548**
F+	0.7968	0.8033	0.8045	**0.8100**
P-	0.8114	0.8229	0.8335	**0.8339**
R-	0.7550	**0.7530**	0.7310	**0.7530**
F-	0.7822	0.7864	0.7789	**0.7914**

MV	BOW	PP	POS	PP & POS
P+	0.7616	**0.7712**	0.7633	0.7668
R+	0.8330	0.8360	**0.8610**	0.8550
F+	0.7957	0.8023	**0.8092**	0.8085
P-	0.8163	0.8210	**0.8406**	0.8362
R-	0.7400	**0.7520**	0.7330	0.7400
F-	0.7763	**0.7851**	0.7831	**0.7850**

DT	BOW	PP	POS	PP & POS
P+	0.7104	0.7241	0.7429	**0.7431**
R+	0.7161	0.7310	0.7295	**0.7450**
F+	0.7133	0.7275	0.7361	**0.7441**
P-	0.7147	0.7284	0.7343	**0.7444**
R-	0.7090	0.7215	**0.7475**	0.7425
F-	0.7118	0.7249	0.7408	**0.7434**

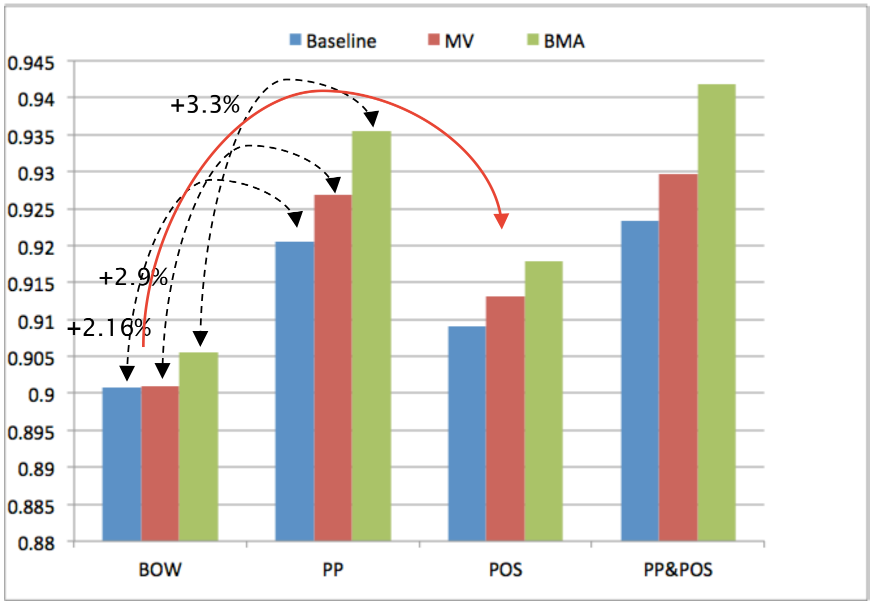
MNB	BOW	PP	POS	PP & POS
P+	**0.7671**	0.7641	0.7635	0.7656
R+	0.8390	0.8455	**0.8650**	0.8605
F+	0.8014	0.8028	**0.8111**	0.8103
P-	0.8229	0.8271	**0.8443**	0.8408
R-	**0.7460**	0.7390	0.7320	0.7365
F-	0.7826	0.7806	0.7841	**0.7852**

SVM	BOW	PP	POS	PP & POS
P+	0.7672	0.7710	0.7765	**0.7800**
R+	0.7568	0.7660	0.7990	**0.8015**
F+	0.7619	0.7685	0.7876	**0.7906**
P-	0.7607	0.7675	0.7930	**0.7959**
R-	0.7710	0.7725	0.7710	**0.7740**
F-	0.7658	0.7700	0.7813	**0.7848**

BN	BOW	PP	POS	PP & POS
P+	0.6692	**0.6965**	0.6824	0.6877
R+	0.7588	0.7150	0.7970	**0.8015**
F+	0.7112	0.7057	0.7352	**0.7402**
P-	0.7224	0.7072	0.7560	**0.7621**
R-	0.6260	**0.6885**	0.6290	0.6450
F-	0.6708	0.6977	0.6867	**0.6987**

- PP negatively contribute to MNB (best baseline) and BN
    - They provide a **small increment** on poor models (SVM and DT); this impact is reflected by MV and BMA
- POS are more informative to identify sarcastic messages
    - PP&POS are the most discriminative features **mostly driven by POS**

RESULTS ON **IRONY**: LINGUISTIC FEATURES



BMA				
	BOW	PP	POS	PP & POS
P+	**0.9218**	0.8988	0.8409	0.8986
R+	0.6900	0.7766	0.7282	**0.7816**
F+	0.7927	0.8333	0.7805	**0.8360**
P-	0.9061	0.9288	0.9133	**0.9302**
R-	0.9700	0.9709	0.9541	**0.9800**
F-	0.9418	0.9493	0.9332	**0.9500**

MV				
	BOW	PP	POS	PP & POS
P+	0.7721	**0.8190**	0.5727	0.7812
R+	0.7286	**0.7842**	0.7392	0.7787
F+	0.7498	**0.8012**	0.6454	0.7799
P-	0.9112	**0.9291**	0.9161	0.9263
R-	0.9283	**0.9422**	0.8377	0.9273
F-	0.9197	**0.9356**	0.8752	0.9268

DT				
	BOW	PP	POS	PP & POS
P+	0.7915	0.8745	0.5710	**0.8714**
R+	0.6544	0.7967	0.6612	**0.7989**
F+	0.7165	**0.8338**	0.6128	0.8335
P-	0.8911	**0.9348**	0.8944	0.9342
R-	0.9425	**0.9619**	0.8525	0.9607
F-	0.9161	**0.9478**	0.8729	0.9476

NB				
	BOW	PP	POS	PP & POS
P+	**0.6199**	0.6161	0.4837	0.5882
R+	**0.7157**	0.6975	0.6699	0.6718
F+	**0.6644**	0.6543	0.5618	0.6273
P-	**0.9001**	0.8945	0.8866	0.8852
R-	**0.8538**	0.8552	0.7831	0.8433
F-	**0.8763**	0.8744	0.8317	0.8637

SVM				
	BOW	PP	POS	PP & POS
P+	0.8712	0.8915	0.6423	**0.8959**
R+	0.7076	0.7756	0.7349	**0.7804**
F+	0.7809	0.8295	0.6855	**0.8341**
P-	0.9083	0.9283	0.9197	**0.9298**
R-	0.9651	0.9685	0.8812	**0.9698**
F-	0.9359	0.9480	0.9000	**0.9494**

BN				
	BOW	PP	POS	PP & POS
P+	0.7680	**0.7992**	0.5727	0.7635
R+	0.7522	**0.8273**	0.7630	0.8017
F+	0.7600	**0.8130**	0.6543	0.7821
P-	0.9180	**0.9417**	0.9231	0.9328
R-	0.9243	**0.9307**	0.8333	0.9172
F-	0.9211	**0.9362**	0.8759	0.9249

- POS tags seem to be important, but they do not characterize irony or non-irony
  - However, they **negatively** contribute to most of the baseline classifiers
- PP are more informative to identify both ironic and non-ironic messages
  - PP&POS are the most discriminative features **mostly driven by PP**

RESULTS ON SARCASM AND IRONY: **MACHINE LEARNING**

- Baseline Classifiers and MV:
  - If PP (or POS) negatively affect a classifier but POS (or PP) positively contributes, then combining them leads to a decrease with respect to POS (or PP)
    - a good feature is not able to make up for the noisy one
  - Including good classifiers in MV does not ensure better performance
    - all the models enclosed in the ensemble should be independent and show different recognition abilities to reduce the bias-variance trade-off

BMA				
	BOW	PP	POS	PP & POS
P+	**0.9218**	0.8988	0.8409	0.8986
R+	0.6900	0.7766	0.7282	**0.7816**
F+	0.7927	0.8333	0.7805	**0.8360**
P-	0.9061	0.9288	0.9133	**0.9302**
R-	0.9700	0.9709	0.9541	**0.9800**
F-	0.9418	0.9493	0.9332	**0.9500**

MV				
	BOW	PP	POS	PP & POS
P+	0.7721	**0.8190**	0.5727	0.7812
R+	0.7286	**0.7842**	0.7392	0.7787
F+	0.7498	**0.8012**	0.6454	0.7799
P-	0.9112	**0.9291**	0.9161	0.9263
R-	0.9283	**0.9422**	0.8377	0.9273
F-	0.9197	**0.9356**	0.8752	0.9268

- BMA:
  - Interesting results with respect to the literature and w.r.t bias-variance trade-off

	Accuracy	Precision	Recall	F-Measure
Reyes et al. [9]	0.8044	0.6610	0.4470	0.5330
BMA	**0.9417**	**0.7814**	**0.8987**	**0.8359**

	BOW	PP & POS
P+	**0.9218**	0.8986
R+	0.6900	**0.7816**
F+	0.7927	**0.8360**
P-	0.9061	**0.9302**
R-	0.9700	**0.9800**
F-	0.9418	**0.9500**

BMA	BOW	PP	POS	PP&POS
DT, SVM, MNB, BN	0.2370	0.2299	0.2434	0.2172
DT, MNB, BN	0.2573	0.2335	0.2707	0.2432
DT, SVM, BN	$\boxed{\mathbf{0.2330}}$	0.2276	0.2482	$\boxed{\mathbf{0.2090}}$
SVM, MNB, BN	0.2531	0.2376	0.2646	0.2509
DT, SVM, MNB	0.2375	0.2287	0.2431	0.2141
DT, SVM	0.2426	0.2317	0.2468	0.2300
DT, MNB	0.2687	0.2361	0.2789	0.2348
DT, BN	0.2595	$\boxed{\mathbf{0.2259}}$	0.2672	0.2329
SVM, MNB	0.2514	0.2366	0.2487	0.2347
SVM, BN	0.2540	0.2430	$\boxed{\mathbf{0.2428}}$	0.2398
MNB, BN	0.2999	0.2800	0.3250	0.3021

SOME (PARTIAL) CONCLUSIONS

- Linguistic Features:
  - All the features are really discriminative

- Machine Learning:
  - Ensemble methods (BMA) can contribute to this field

RECAP

- Choose the **sentiment paradigm** that is more convenient according to the application domain taking into account the **computational complexity** of each approach:
  - If there is a **lexicon**, use it (or induce it)
  - If you don't have a lexicon, but in your problem you can exploit some distant supervision you can take advantage of **semi-supervised** approaches
  - If your problem is based on common-sense, use **semantic approaches**
  - If your problem is in a closed-domain where you have labelled examples, use **supervised paradigms** (eventually ensemble)
  - If you want to address the problem as **aspect-based** sentiment analysis, follow an unsupervised paradigm.
  - If **irony** (sarcasm) is present, then reverse the polarity!
- Considering some **linguistic features** (chunk, POS, social network language as emoticon, hashtags, etc...) can be useful to improve the prediction ability of sentiment and irony/sarcasm.