

Outline

1. Witty language
 - Humor generation
 - Humor recognition
2. Affective Text
 - Lexical resources
 - Annotation of emotions in text
 - Colors of emotions in texts
 - Dancing with words
3. Persuasive NLP
 - Analyzing political speeches along with audience reactions (e.g. applauses)
 - How to evaluate persuasive language ?
4. Deceptive Language recognition
 - Is it possible to recognize when people are lying, just using the produced text ?

"You should not trust the devil, even if he tells the truth."

– Thomas of Aquin (medieval philosopher)

- From a NLP perspective:
 - Lying is a particular cognitive activity
 - Is it possible that this has some influence in the language used ?

Explorations in the Automatic Recognition of Deceptive Language

- This work presents some initial experiments in the recognition of deceptive language:
 - We introduce *three data sets* of true and lying texts
 - We show that *automatic classification* is a viable technique to discriminate between truth and falsehood, as expressed in language
 - We introduce a method for *class-based feature analysis*, for characterizing some properties of deceptive texts

R. Mihalcea and C. Strapparava

"The Lie Detector: Explorations in the Automatic Recognition of Deceptive Language"
ACL 2009

Motivations

- The discrimination between truth and falsehood seems like the holy grail of human capabilities
- Significant attention from philosophy, psychology and sociology
- The approach: recognition of deceptive language from a *data-driven perspective*, salient features of lying texts using NLP
- We deliberately focus on *written language*, since it represents the type of data most frequently encountered on the **Web** (e.g., chats, forums, etc...) or in other collections of documents

Motivations

- Lying is a peculiar human activity
- It requires ability, cleverness and imagination
- Surely there is a connection with the behavior and in particular there were studies about emotive reaction



- Does lying activity leave some traces in the used language ?

Data Sets

- A corpus with explicit labeling of truth value
- We had to create one ourselves
=> Amazon Mechanical Turk
- Three different topics:
 - Opinions on *abortion*
 - Opinions on *death penalty*
 - Feelings about *best friend*

Data Sets (cont.)

- Instructions to the contributors
- For the topics *abortion* and *death penalty* :
 - Imagine they were taking part in a debate
 - They were asked to prepare a brief speech expressing their *true* opinion and
 - then a second one expressing the opposite (thus *lying*) of their opinion
- For the topic *best friend* :
 - Think about a great friend and describe the reasons for their friendship
 - Think about a person they could not stand, and describe as a best friend
- In all the cases at least 4-5 sentences

Data Sets (cont.)

- For each topic we collected 100 true and 100 false statements
- Average of 85 words per statement
- Manual verification of the quality of the contributions
- With two exceptions, the entries were found of good quality

Data Sets: Sample Entries

TRUTH	LIE
ABORTION	
I believe abortion is not an option. Once a life has been conceived, it is precious. No one has the right to decide to end it. Life begins at conception, because without conception, there is no life.	A woman has free will and free choice over what goes on in her body. If the child has not been born, it is under her control. Often the circumstances an unwanted child is born into are worse than death. The mother has the responsibility to choose the best course for her child.
DEATH PENALTY	
I stand against death penalty. It is pompous of anyone to think that they have the right to take life. No court of law can eliminate all possibilities of doubt. Also, some circumstances may have pushed a person to commit a crime that would otherwise merit severe punishment.	Death penalty is very important as a deterrent against crime. We live in a society, not as individuals. This imposes some restrictions on our actions. If a person doesn't adhere to these restrictions, he or she forfeits her life. Why should taxpayers' money be spent on feeding murderers?
BEST FRIEND	
I have been best friends with Jessica for about seven years now. She has always been there to help me out. She was even in the delivery room with me when I had my daughter. She was also one of the Bridesmaids in my wedding. She lives six hours away, but if we need each other we'll make the drive without even thinking.	I have been friends with Pam for almost four years now. She's the sweetest person I know. Whenever we need help she's always there to lend a hand. She always has a kind word to say and has a warm heart. She is my inspiration.

Experiments

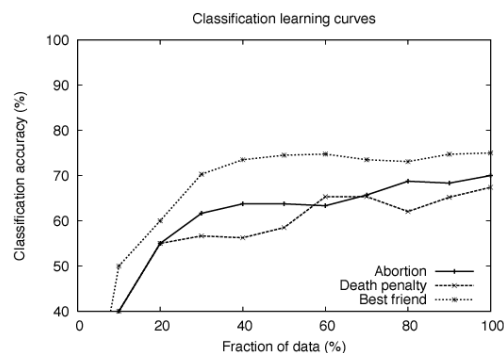
- Two classifiers: Naïve Bayes (NB) and SVM
- Minimal preprocessing: tokenization, stemming, no feature selection, no stopword removal
- Ten-fold cross validation results:

Topic	NB	SVM
<i>Abortion</i>	70.0%	67.5%
<i>Death Penalty</i>	67.4%	65.9%
<i>Best Friend</i>	75.0%	77.0%
<i>Average</i>	70.8%	70.1%

Baseline 50%

Learning curves

- The overall growing trend indicates that more data is likely to improve the accuracy
- Collecting additional data as a possible step for future work



Cross-topic classification

- Testing the portability of the classifiers among topics
- Accuracy is still significantly higher than the baseline
- The learning process is not bound to a specific topic

Training	Test	NB	SVM
<i>Death Penalty + Best Friend</i>	<i>Abortion</i>	62.0%	61.0%
<i>Abortion + Best Friend</i>	<i>Death Penalty</i>	58.7%	58.7%
<i>Abortion + Death Penalty</i>	<i>Best Friend</i>	58.7%	53.6%
<i>Average</i>		59.9%	57.8%

Dominant Word Classes

- A measure of saliency for a given word class inside a collection of deceptive/truthful texts
- Given a word class $C = \{w_1, w_2, \dots, w_n\}$, its coverage with respect to deceptive (D) and truthful (T) corpora respectively

$$Coverage_D(C) = \frac{\sum_{w_i \in C} Frequency_D(w_i)}{Size_D}$$

$$Coverage_T(C) = \frac{\sum_{w_i \in C} Frequency_T(w_i)}{Size_T}$$

- where $Frequency_D(w_i)$ represents the occurrences of w_i inside D
- The **dominance score** of the class C in the deceptive corpus wrt. the truthful one is

$$Dominance_D(C) = \frac{Coverage_D(C)}{Coverage_T(C)}$$

Word Classes

- We exploit the word classes as defined in *Linguistic Inquiry and Word Count* (LIWC) resource
- LIWC (Pennbaker and Francis) was developed as a resource for psycholinguistic analysis
- 2,200 words and word stems grouped into about 70 broad categories (e.g., *emotion*, *cognition*, ...)
- Validated in many psycholinguistic studies

Dominant Word Classes

- Both in deceptive and truthful texts:
 - Three of the top five dominant classes are related to humans
- In **deceptive** texts:
 - The human word classes (*You, Other, Humans*) represent a detachment from the self
 - Words related to certainty (*Certain*)
- In **truthful** texts:
 - The human word classes (*I, Friends, Self*) are closely connected to the self
 - Belief-oriented vocabulary (*Insight*) including words such as believe, feel, think

Dominant Word Classes

- Some examples of dominant word classes in deceptive texts

Class	Score	Sample words
Deceptive Text		
METAPH	1.71	god, die, sacred, mercy, sin, dead, hell, soul, lord, sins
YOU	1.53	you, thou
OTHER	1.47	she, her, they, his, them, him, herself, himself, themselves
HUMANS	1.31	person, child, human, baby, man, girl, humans, individual, male, person, adult
CERTAIN	1.24	always, all, very, truly, completely, totally
Truthful Text		
OPTIM	0.57	best, ready, hope, accepts, accept, determined, accepted, won, super
I	0.59	I, myself, mine
FRIENDS	0.63	friend, companion, body
SELF	0.64	our, myself, mine, ours
INSIGHT	0.65	believe, think, know, see, understand, found, thought, feels, admit

Conclusions and Future Work

- Automatic recognition of deceptive language in written texts:
 - Three data sets on different topics
 - Truthful and lying texts are separable
 - Analysis of classes of salient features, with insights into the vocabulary used in deceptive texts
- From the learning curves, more data is likely to improve the accuracy => Collection of additional data
- Exploring the role of affect and its possible integration in the recognition of deceptive language

Outline

1. Witty language
 - Humor generation
 - Humor recognition
2. Affective Text
 - Lexical resources
 - Annotation of emotions in text
 - Colors of emotions in texts
 - Dancing with words
3. Persuasive NLP
 - Analyzing political speeches along with audience reactions (e.g. applauses)
 - How to evaluate persuasive language ?
4. Deceptive Language recognition
 - Is it possible to recognize when people are lying, just using the produced text ?