NLP meets Computational Social (Media) Science

Dirk Hovy
Computer Science Department
University of Copenhagen

mail@dirkhovy.com
@dirk_hovy
The only other Republican governor in a competitive re-election race this year, according to Cook Political Report’s ratings, is Pat McCrory of North Carolina, who hasn’t made a presidential endorsement. However, there are 11 Republican senators and 34 Republican members of the House who face competitive races, according to Cook. The only one to have endorsed Trump is Tom Reed, the incumbent from New York’s 23rd Congressional District, a Republican-leaning swing district that covers much of the rural, western part of the state.

Understanding supply and demand is easy. What is difficult to comprehend is what makes people like a particular stock and dislike another stock. This comes down to figuring out what news is positive for a company and what news is negative. There are many answers to this problem and just about any investor you ask has their own ideas and strategies.


Follow us: Investopedia on Facebook
Natural Language Processing

Linguistics

Computer Science

informed linguistic hypotheses
large-scale statistical analysis
I admire Rosa Parks.
Who buys what?

"min søn"

BUSINESS
Who buys what?

"engagement ring"

BUSINESS
### Sexism in Politics

#### Tone
- **Positive**
- **Negative**
- **Neutral**

<table>
<thead>
<tr>
<th>Tone</th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Neutral</td>
</tr>
<tr>
<td>Negative</td>
<td>Negative</td>
<td>Negative</td>
<td>Neutral</td>
</tr>
<tr>
<td>Neutral</td>
<td>Neutral</td>
<td>Neutral</td>
<td>Neutral</td>
</tr>
</tbody>
</table>

#### Key Words
- **b*ch**
- **victims**
- **loss**
- **losing**
- **bye**
- **lgbt**
- **POL. SCI.**

#### Other Words
- **benghazi**
- **husband**
- **fbi**
- **injustice**
- **angry**
- **unborn**
- **shame**
- **liars**
- **queen**
- **killary**
- **goldman**
- **email**
- **woman**
- **advice**
- **women**
- **flint**
- **ways**
- **maker**
- **law**
- **affect**
- **remarks**
- **human**
- **dm**
- **shes**
- **blame**
- **lady**
- **children**
- **advice**
- **wing**
- **pol. sci.**

#### URLs
- [Link to article](https://www.washingtonpost.com/news/monkey-cage/wp/2016/02/24/these-6-charts-show-how-much-sexism-hillary-clinton-faces-on-twitter/)
Censorship

POL. SCI.
Q: What area of London should I live in?

A: Cool areas to live in at the moment are: / Clapham / Balham / Battersea / Hoxton / Camden
Language Change

Sociolinguistics
Measuring Language Variation

**Sociolinguistics**

Age and gender

Gender distribution

Age distribution

Total records with NUTS information: 1,048

The map shows the percentage of records for each NUTS region that use the term træls.

The term is used more frequently than average in following regions: Byen København (DK011), Fyn (DK031), Nordjylland (DK050), Sydjylland (DK032), Vestjylland (DK041), Østjylland (DK042) (the ratio of term in the vocabulary of these regions is larger than the average ratio across regions).
Visualizing Language Variation

SOCIOLINGUISTICS
Personality

introverted-extroverted

thinking-feeling

Plank & Hovy (EMNLP 2015)
The Transparent User

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>Rosenthal and McKeown, 2011; Nguyen et al., 2011</td>
</tr>
<tr>
<td>gender</td>
<td>Alowibdi et al., 2013; Ciot et al., 2013; Liu and Ruths, 2013; Volkova et al.</td>
</tr>
<tr>
<td>profession</td>
<td>Preoțiuc-Pietro et al., 2015</td>
</tr>
<tr>
<td>income</td>
<td>Preoțiuc-Pietro et al., 2015b</td>
</tr>
<tr>
<td>personality</td>
<td>Plank and Hovy, 2015</td>
</tr>
<tr>
<td>political views</td>
<td>Volkova et al., 2014</td>
</tr>
<tr>
<td>location</td>
<td>Bergsma et al., 2013; Rahimi et al., 2016</td>
</tr>
</tbody>
</table>
Workshop on NLP and Computational Social Sciences

@ACL
Aug 3, 2017
Vancouver, Canada

https://sites.google.com/site/nlpandcss/

@nlpandcss
That’s it, NLP can help social science and business

All is great!
...no, really...
Ok, so maybe there’s one thing
The Assumption

All language is i.i.d.
Language Distributions

age

income

WSJ
Shannon Game

What's the next word?

The
A
Then
If
When
My
He
You
I
...

house
friend
dog
car
water
hovercraft
pants
God
word
...

with
water

eels

...?

...?

How much longer?
Shannon Game

Very surprising

\[- \sum_{x} p(x) \log p(x)\]

Information

TOTALLY PREDICTABLE

TOTALLY PREDICTABLE

0
Language

- Information: 35%
- Representation: 29%
- Gossip: 11%
- Strategy: 10%
- Support: 8%
- Jokes: 7%
Demographics

- Depression
- under 35
- low income
- female
- college education

F0 > 180Hz

Nguyen et al., 2011; Alowibdi et al., 2013; Ciot et al., 2013; Liu and Ruths, 2013; Volkova et al., 2014; Volkova et al., 2015; Plank and Hovy, 2015; Preotiuc-Pietro et al., 2015a/b
The Problem

NLP performance correlates with demographics.

Not definitely sure yet

distance from “standard”
The Consequences

NLP WORKS WELL

population
More Data?

**How can we add extra-linguistic info to NLP models?**
Improving NLP with Demographic Information
Example 1: Text Classification

This is a tiny little example text written by someone.

This is a tiny little example text written by someone.

This is a tiny little example text written by someone.

This is a tiny little example text written by someone.
### Data

<table>
<thead>
<tr>
<th>task</th>
<th>labels</th>
<th>train</th>
<th>test</th>
</tr>
</thead>
<tbody>
<tr>
<td>topic classification</td>
<td>5</td>
<td>739k</td>
<td>493k</td>
</tr>
<tr>
<td></td>
<td></td>
<td>126k</td>
<td>84k</td>
</tr>
<tr>
<td>sentiment analysis</td>
<td>3</td>
<td>345k</td>
<td>230k</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72k</td>
<td>48k</td>
</tr>
<tr>
<td>age/gender classification</td>
<td>2</td>
<td>301k</td>
<td>201k</td>
</tr>
<tr>
<td></td>
<td></td>
<td>301k</td>
<td>201k</td>
</tr>
</tbody>
</table>

, 2 genders, 2 age groups

Hovy (2015)
Embeddings
Systems

**AGNOSTIC**

+ training data

**INFORMED**

+ training data

choose

Hovy (2015)
Results for Age (avg)

Hovy (2015)
Results for Gender (avg)

- **sentiment**: Agnostic - 61.93, Awared - 62.92
- **topic**: Agnostic - 55.95, Aware - 57.22
- **age classification**: Agnostic - 59.19, Aware - 59.95

*Hovy (2015)*
Example 2:

Multitask Learning for Mental Health Conditions
Mental Health Risk Prediction

NLP

suicide?

suicidal ideation

Psychologists

stigma
availability
accuracy/FN
Comorbidity and Correlation

- age
- gender
- PTSD
- depression
- ...

**Comorbidity and Correlation**

- Age and gender correlate with PTSD and depression.
- Depression and PTSD are also correlated.
Motivation from “The Karate Kid”

Single-task Learners (STL)

Multitask Learner (MTL)
## Data

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurotypicality</td>
<td>4791</td>
</tr>
<tr>
<td>Anxiety</td>
<td>2407</td>
</tr>
<tr>
<td>Depression</td>
<td>1400</td>
</tr>
<tr>
<td>Suicide attempt</td>
<td>1208</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>749</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>349</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>263</td>
</tr>
<tr>
<td>PTSD</td>
<td>248</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>191</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td><strong>9611</strong></td>
</tr>
</tbody>
</table>
Results ROC

Suicide Attempt

Non-Neurotypicality
Results: TPR@FPR=0.1

~120 AT-RISK PATIENTS

NNT
Anxiety
Depression
Suicide
Eating
Schizophrenia
Panic
PTSD
Bipolar

Logistic Regression
Feed Forward
MTL no gender
MTL w/ gender
So then all is good, right?
...right?
Challenges
Ethics and NLP

35  male  German  college education

Exclusion  Overgeneralization  Exposure  Dual Use
Exclusion

street -> skreet
brother -> brotha

avg. LL for non-AAVE Tweets only

78.57%
Exclusion

Accuracy

POS-tagging

600 user reviews

CoNLL trained

Hovy/Søgaard (ACL 2015)
Overgeneralization

FALSE POSITIVES

Aug 6 2016

Dear Ms. Hovy,

Congratulations on reaching retirement age!

Also, you’re on a no-fly list because of your political views and religious beliefs.
The Signal and the Noise

Saeidi et al., 2017

violent 0.44
gang 0.43
drug 0.42
rob 0.4
danger 0.39
knife 0.39
integration 0.38
black 0.38
boy 0.38
evenly 0.38
dangerous 0.37
stab 0.37
Hate Speech Detection

Abusive Language Workshop
ACL Vancouver, Aug 4 2017
Exposure

Underexposure:
- Available: 65%
- Not available: 35%
- Treebanks

Overexposure:
- New York City: 8.5m
- Lagos: 16m
- Sentiment analysis
- Discourse parsing
- Bias

Evaluation:
- 87%
- Semantic resources
## Dual Use

<table>
<thead>
<tr>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>authorship attribution</strong></td>
<td>historical documents</td>
</tr>
<tr>
<td><strong>text classification</strong></td>
<td>sentiment analysis</td>
</tr>
<tr>
<td><strong>personalization</strong></td>
<td>better user experience</td>
</tr>
<tr>
<td></td>
<td>dissenter anonymity</td>
</tr>
<tr>
<td></td>
<td>censorship</td>
</tr>
<tr>
<td></td>
<td>tailored ads</td>
</tr>
</tbody>
</table>

Hovy/Spruit, 2016
Dear Ms. Hovy,

Congratulations on reaching retirement age!

Also, you’re on a no-fly list because of your political views and religious beliefs.

What can we do?

<table>
<thead>
<tr>
<th>Problem</th>
<th>Source</th>
<th>Countermeasures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overfitting</td>
<td>data selection</td>
<td>regularization, priors, sampling</td>
</tr>
<tr>
<td>Overgeneralization</td>
<td>models</td>
<td>dummy labels, error weighting, confidence thresholds</td>
</tr>
<tr>
<td>Exposure</td>
<td>research design</td>
<td>consider possible impact</td>
</tr>
<tr>
<td>Dual Use</td>
<td>community goals</td>
<td>educate users, keep discussion going</td>
</tr>
</tbody>
</table>

Hovy/Spruit, 2016
Slow internet, fast access version is here: http://www.ethicsinnlp.org/workshop/program.html

Twitter Hashtag: #EthNLP

Ethics in Natural Language Processing

A Workshop at EACL 2017

4 April 2017
Valencia, Spain

Previous Updates

http://www.ethicsinnlp.org/
Wrapping up
Conclusion

• social media + NLP = new methods for CSS

• but: NLP is more than engineering

  • language \geq information, reflects demographics

  • accounting for demographics affects performance

• Work on social media has social impact and responsibility
Open questions

- how do we reflect ever-changing language?
- how do we account for demographic variation?
- how do we ensure fairness?
Thanks!
Questions?

www.dirkhovy.com
@dirk_hovy
Workshop on NLP and Computational Social Sciences

@ACL
Aug 3, 2017
Vancouver, Canada

https://sites.google.com/site/nlpandcss/

@nlpandcss