University of Stuttgart
Institute for
Natural Language Processing

SEMALYTIX

Emotion Analysis

between Academia, Industry, Linguistics, Humanities, and Computer Science

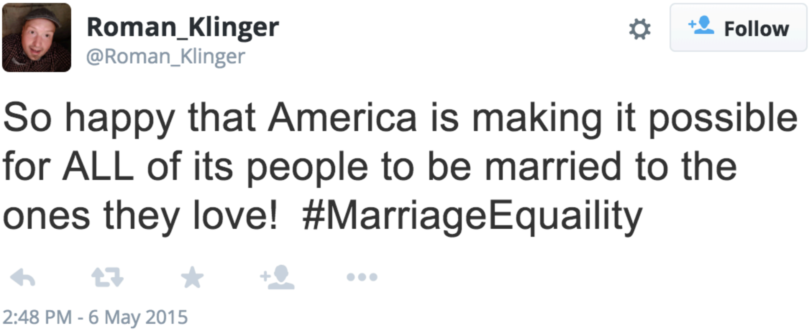
AI2Future, Zagreb, Croatia

October 11/12, 2018

Roman Klinger

@roman_klinger romanklinger
http://www.romanklinger.de/

Motivation



Which emotions are expressed?

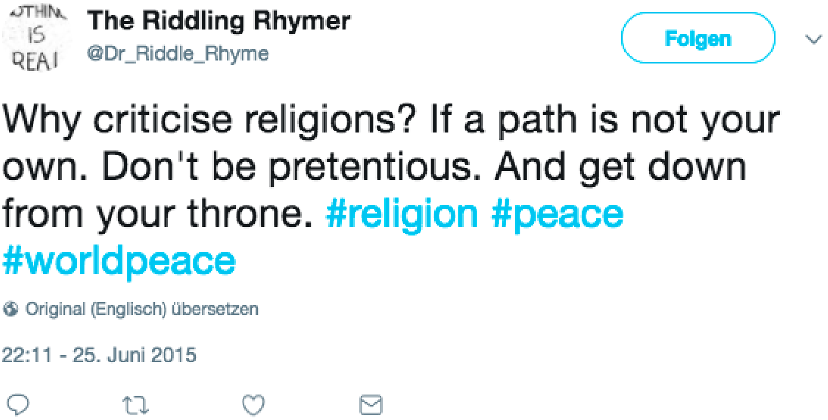
Anger Anticipation Disgust Fear Joy Sadness Surprise Trust

Motivation

Which emotions are expressed?

Anger Anticipation Disgust Fear Joy Sadness Surprise Trust

Motivation



Which emotions are expressed?

Anger Anticipation Disgust Fear Joy Sadness Surprise Trust

Goal of this Presentation

- Emotion analysis is not just a more fine-grained version of sentiment analysis.
- Challenging task.
- Many open research and application questions.

Outline

What is Emotion Analysis?

Sentiment analysis

positive vs. negative
(neutral, mixed)

Subjectivity analysis

subjective vs. objective

Emotion analysis discrete (Ekman/Plutchik)

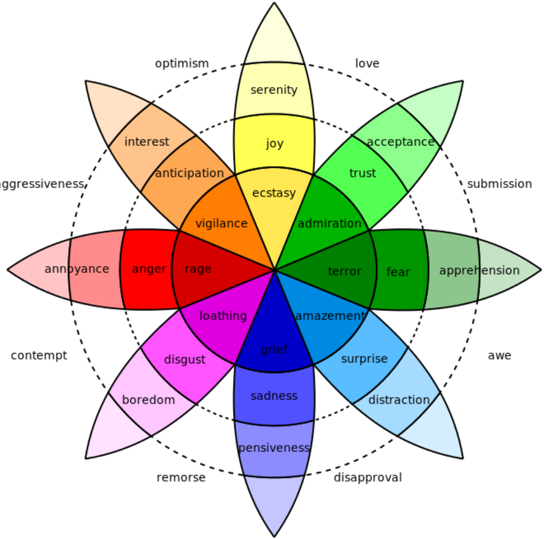
discrete emotion classes

Emotion analysis cont. (Posner/Russell/Peterson)

valence and arousal

Emotion Models: Ekman

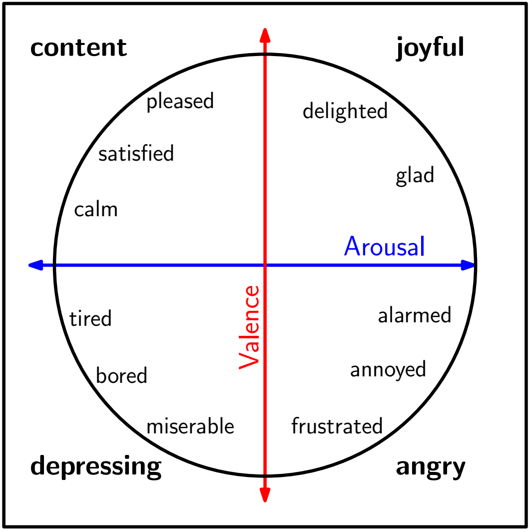
Emotion Models: Plutchik's Wheel



- Anger
- Anticipation
- Disgust
- Fear
- Joy
- Sadness
- Surprise
- Trust

Models opposing emotions and intensity!

Emotion Models: Continuous



What are emotions? Physiological reactions?

James-Lange Theory (1884, 1885)

$\Longrightarrow$

Heart pounding, trembling, sweating, running away

$\Longrightarrow$

Cannon-Bard Theory ($\approx$ 1925)

Emotions are independent of physiological signals

- backed by experiments: no reliable correlation between physiological changes and emotions

$\Rightarrow$ physiological reaction and emotion reaction are independent

What are emotions? Physiological interactions?

Valins Effect (Stuart Valins, 1966)

Appraisal Theories: Why are there emotions?

Scherer, 2005

Emotions are "an episode of interrelated, synchronized changes ... in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism"

Summary

Emotions have different components...

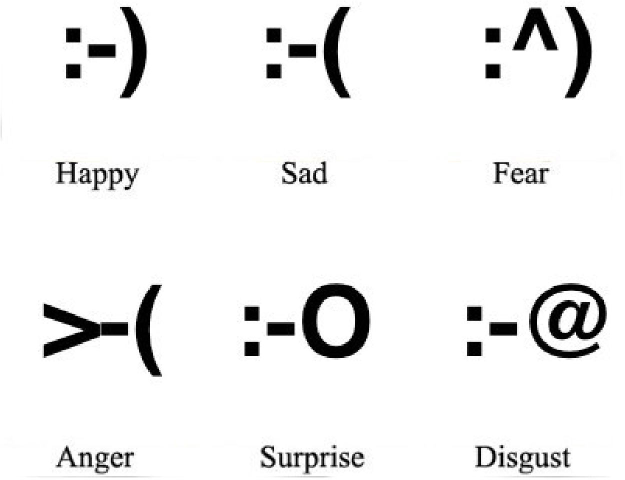
- Cognitive appraisal
- Bodily symptoms
- Reactions
- Expression
- Subjective perceptions
- ...

...what we do with text?

Outline

1. Introduction: What is Emotion Analysis?
2. Emotion Classification in Social Media
3. Emotion Classification in the Humanities
4. Emotion Classification for the Pharma Industry
5. Conclusion

Ekman in Social Media



Corpora

Dataset	Type	Annotation	Size	Source	Avail.
AffectiveText		+ {valence}	1,250	Strapparava (2007)	D-U
Blogs		+ {mixed, noemo}	5,025	Aman (2007)	R
CrowdFlower		+ {fun, love, …}	40,000	Crowdflower (2016)	D-U
DailyDialogs			13,118	Li et al. (2017)	D-RO
Electoral-Tweets			4,058	Mohammad (2015)	D-RO
EmoBank			10,548	Buechel (2017)	CC-by4
EmoInt		- {disgust, surprise}	7,097	Mohammad (2017)	D-RO
Emotion-Stimulus		+ {shame}	2,414	Ghazi et al. (2015)	D-U
fb-valence-arousal			2,895	Preoţiuc (2016)	D-U
Grounded-Emotions		☺☹	2,585	Liu et al. (2017)	D-U
ISEAR		+ {shame, guilt}	7,665	Scherer (1997)	GPLv3
Tales			15,302	Alm et al. (2005)	GPLv3
SSEC			4,868	Schuff et al. (2017)	D-RO
TEC		+ {±surprise}	21,051	Mohammad (2012)	D-RO

Bostan/Klinger, COLING 2018

Task Description and Research Question

Corpus Generation Task

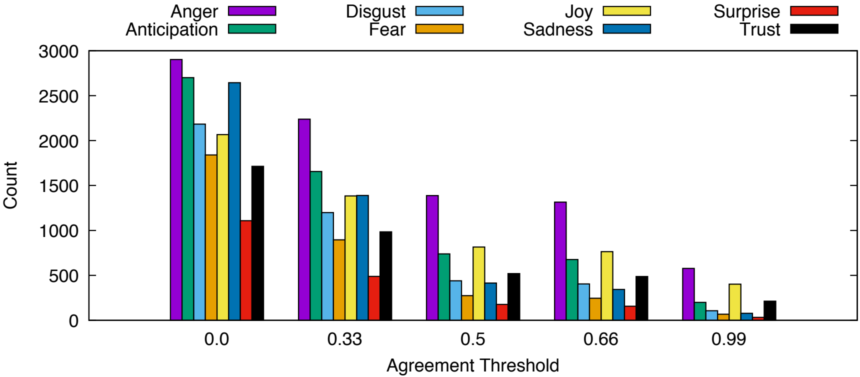
- 4870 Tweets with preexisting annotation of sentiment and stance (SemEval 2016)

Research Questions

- What's the inter-annotator agreement?
- Which annotation layers interact?
- How well is it possible to computationally estimate such annotations?

Schuff et al, WASSA 2017

Label Counts



⇒ Seldom that all annotators agree
⇒ Low number of majority vote annotations
⇒ Low quality of annotation combination?

Difficult Examples (1)

Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
> 0.33					> 0.33	> 0.33	

Difficult Examples (2)

"2 pretty sisters are dancing with cancered kid"

Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
			> 0.0	> 0.0	> 0.0		> 0.0

Cooccurrences of Labels

	Emotions								Sentiment			Stance		
	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Positive	Negative	Neutral	In Favor	Against	None
Anger	**2902**	1437	1983	1339	774	2065	711	640	275	2534	93	630	1628	644
Anticipation		**2700**	1016	1029	1330	1369	482	1234	1094	1445	161	772	1291	637
Disgust			**2183**	1024	512	1628	526	404	126	2008	49	429	1291	463
Fear				**1840**	466	1445	407	497	306	1445	89	448	982	410
Joy					**2067**	682	438	1101	1206	750	111	596	952	519
Sadness						**2644**	664	613	345	2171	128	604	1429	611
Surprise							**1108**	222	219	801	88	257	521	330
Trust								**1713**	1082	558	73	500	860	353
Positive									**1524**	0	0	485	673	366
Negative										**3032**	0	622	1665	745
Neutral											**312**	97	71	144
In Favor												**1204**	0	0
Against													**2409**	0
None														**1255**

- Many cooccurrences as expected (pos w/ pos, neg w/ neg)Positive Anger Negative Joy Positive Disgust

Examples

Positive Anger

"Lets take back our country! Whos with me? No more Democrats!2016"
"Why criticise religions? If a path is not your own. Don't be pretentious. And get down from your throne."

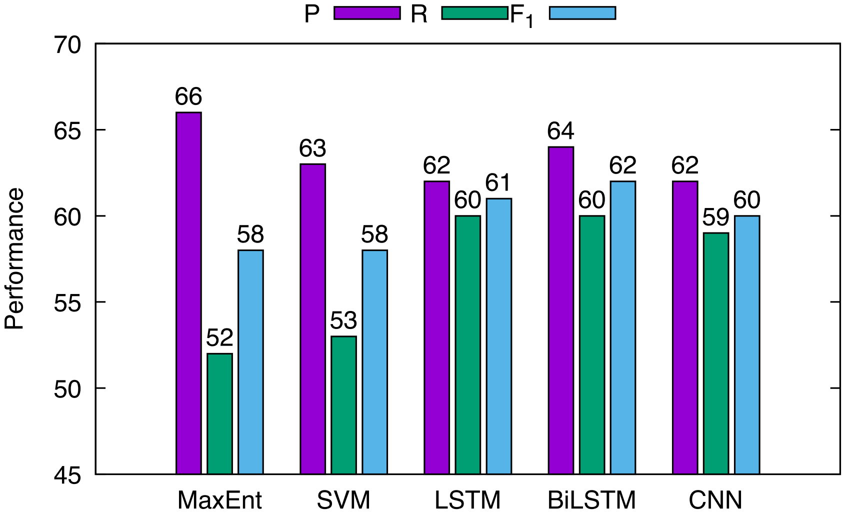
Negative Joy

"Global Warming! Global Warming! Global Warming! Oh wait, it's summer."
"I love the smell of Hillary in the morning. It smells like Republican Victory."

Positive Disgust

"#WeNeedFeminism because #NoMeansNo it doesnt mean yes, it doesnt mean try harder!"

Models


P
R
F_1
70
65
60
55
50
45
Performance
66
52
58
63
53
58
62
60
61
64
60
62
62
59
60
MaxEnt
SVM
LSTM
BiLSTM
CNN

Conclusion: Emotion Classification in Social Media

- Deep learning better than traditional models (but super-slow)
- Still difficult problem with challenging instances
- Enables a lot of analysis tasks: Query for tweets including specific entities to evaluate the associated emotion (we will come back to that later)

Outline

1. Introduction: What is Emotion Analysis?
2. Emotion Classification in Social Media
3. Emotion Classification in the Humanities
4. Emotion Classification for the Pharma Industry
5. Conclusion

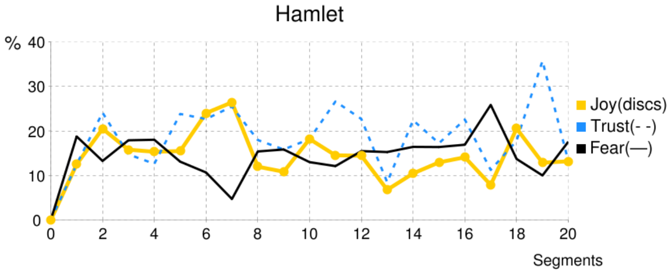
Literary Studies

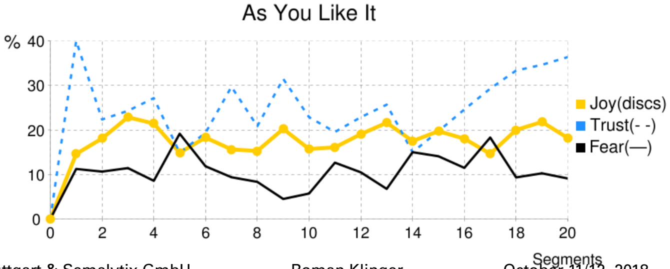
Our Research Question

Can we characterize literature with the help of emotion analysis?

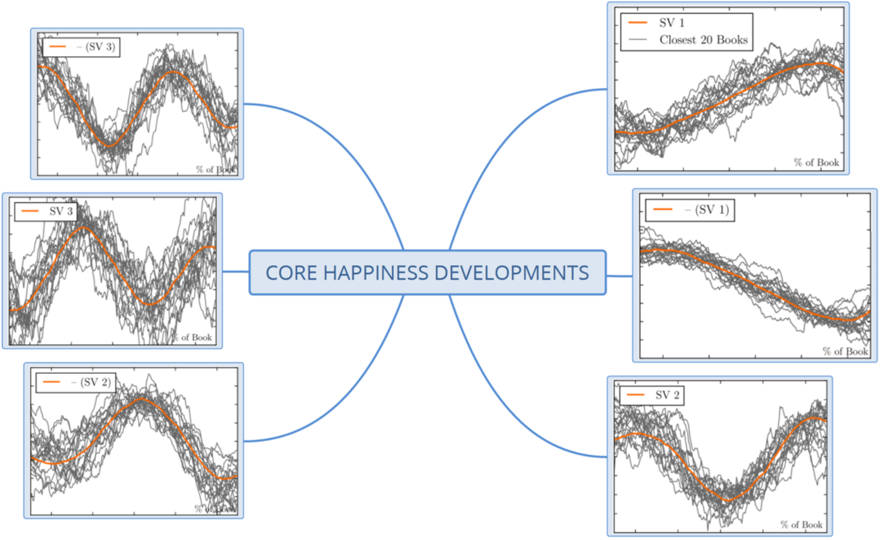
Gaspar Melchor de Jovellanos
painted by
Francisco José de Goya y Lucientes

Previous Work: Mohammad, 2011





Previous Work: Reagan, 2016



Can we use this information to predict genres?

Setting

Kim et al, LaTeCH-CLfL 2017

Setting

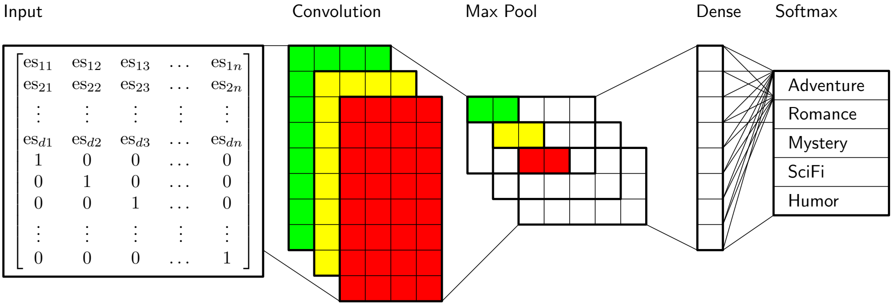
Features

- Emotion scores
- Bag-of-words
- Emotion words

Models

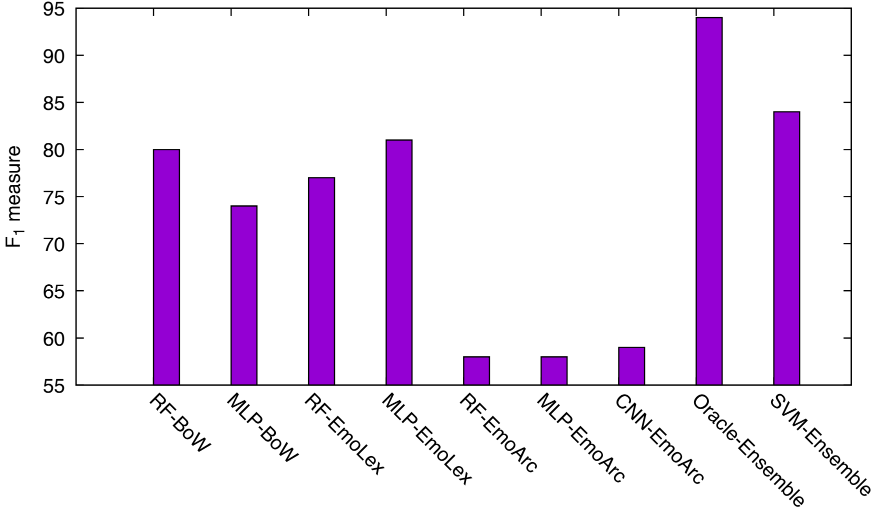
- Random Forrest, Multi-layer Perceptron, CNN

CNN

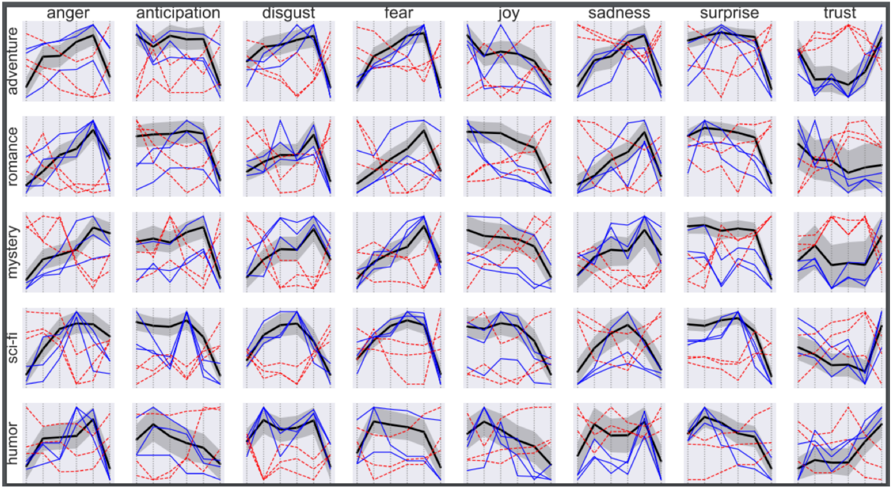


Input
Convolution
Max Pool
Dense
Softmax
es11 es12 es13 ... es1n
es21 es22 es23 ... es2n
esd1 esd2 esd3 ... esdn
1 0 0 ... 0
0 1 0 ... 0
0 0 1 ... 0
0 0 0 ... 1
Adventure
Romance
Mystery
SciFi
Humor

Results for Genre Classification



Genres and Emotion

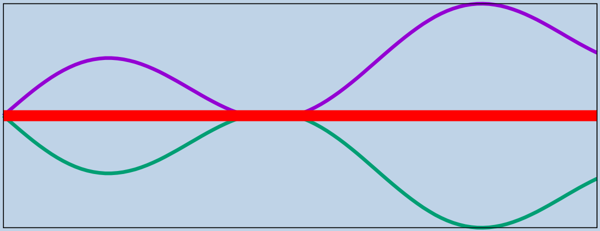


Summary and Future

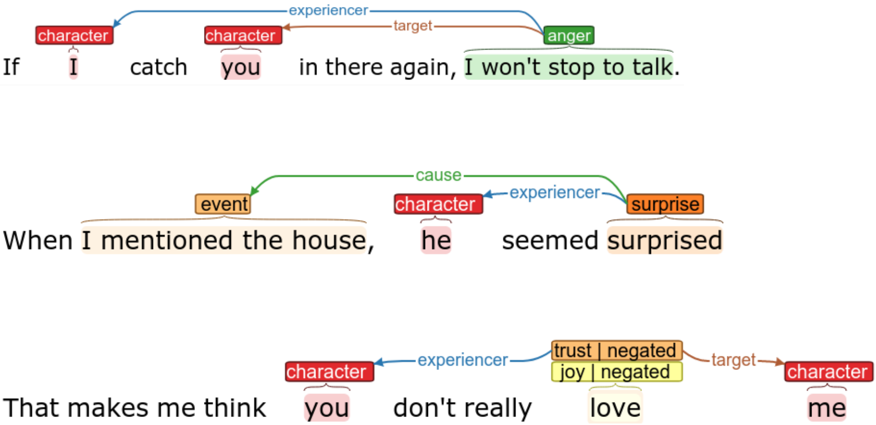
- Emotion curves and words predict genres clearly better than random

Challenge

Approach ignores differences in characters and interactions

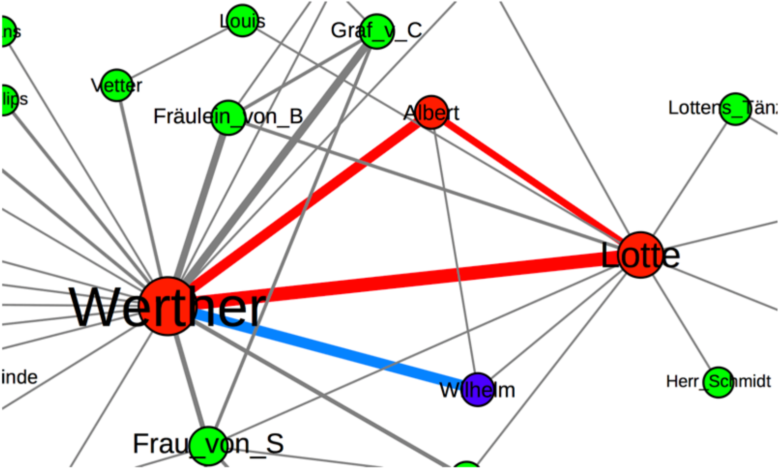


Current Work: Structured Emotion Prediction



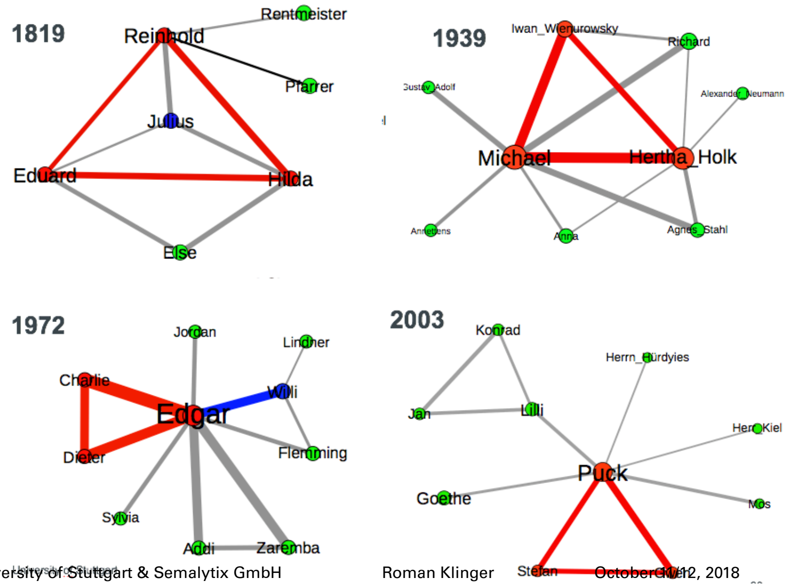
REMAN Corpus (Kim, COLING 2018)

Emotion Character Networks



(Barth, Kim, Murr, Klinger. DHd 2018)

Emotion Character Networks



IEST

Implicit Emotions Shared Task

Example Instances

- It's EMOTION when you feel like you are invisible to others.
- My step mom got so EMOTION when she came home from work and saw that the boys didn't come to Austin with me.

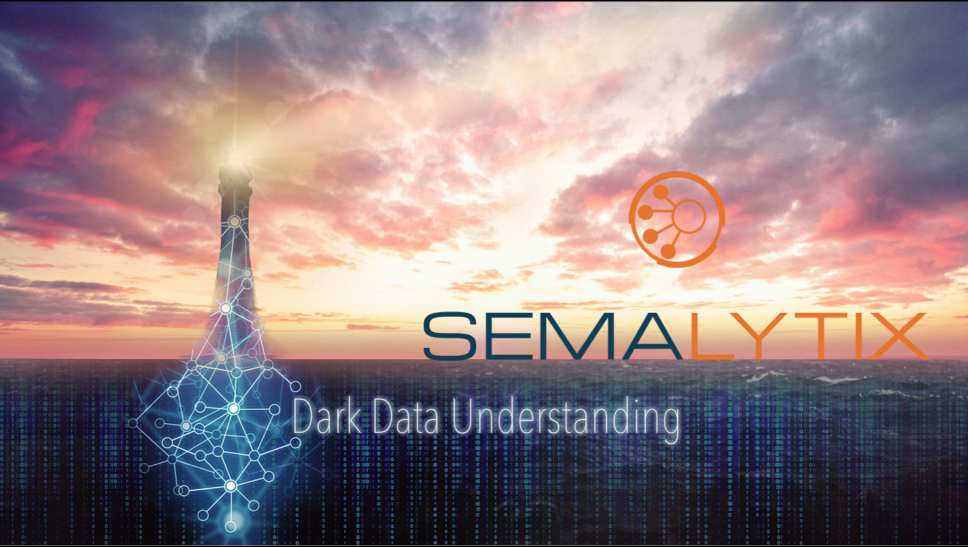
Material online

`http://implicitemotions.wassa2018.com`

Klinger, De Clercq, Mohammad, Balahur, WASSA 2018

Outline

1. Introduction: What is Emotion Analysis?
2. Emotion Classification in Social Media
3. Emotion Classification in the Humanities
4. Emotion Classification for the Pharma Industry
5. Conclusion


SEMALYTIX
Dark Data Understanding

BEACON™ scales Dark Data Understanding by combining:

- cutting-edge analytics
- data visualisation
- full service integration

It takes unstructured text and structures it. It understands the text’s context the way a human expert would. It pro-vides highly intuitive and actionable insights, in real-time.

Problem

Commercialisation

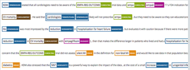
Systematically and reliably hearing the authentic patient voice on social and medical platforms in order to generate insights toward market access **in real time.**

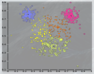
forums
blogs
Reddit/Twitter
unstructured text

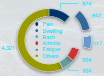
Building internal data science and customer value divisions are
expensive and/or inaccessible

"I imagine mum would have also have been suffering something similar as well as her other symptoms. She was put on antipsychotics. Sadly though, she is still suffering delusions to some degree."

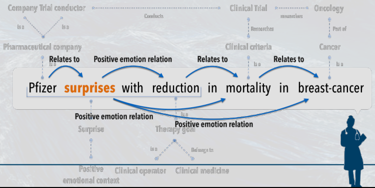
Data turned into intuitive dashboards

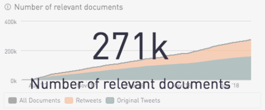




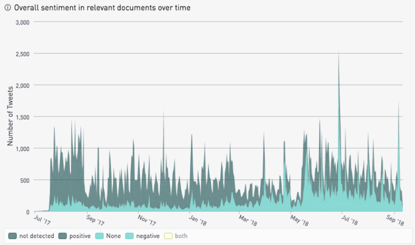


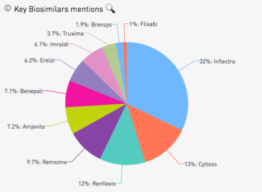
"…have found the treatment easy to deal with […] I've suffered minor things such as higher than normal temperatures, joint aches, itching, appetite loss and joint aching […] none of these side effects have stopped me operating as normal tho."

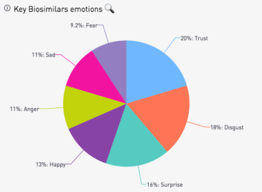


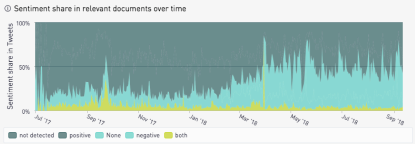


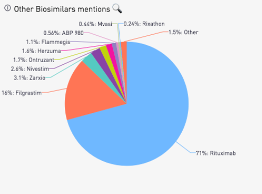
Documents crawled between 2017-07-01 and 2018-09-12

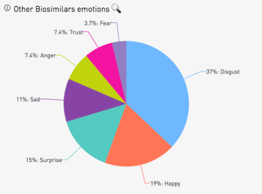












Examples

Anger

If you suffer from psoriasis, postnatally prepare to turn into a hideous scaly beast! #ShitTheyDontTellYou #LongSleevesAgain #ItchyAF

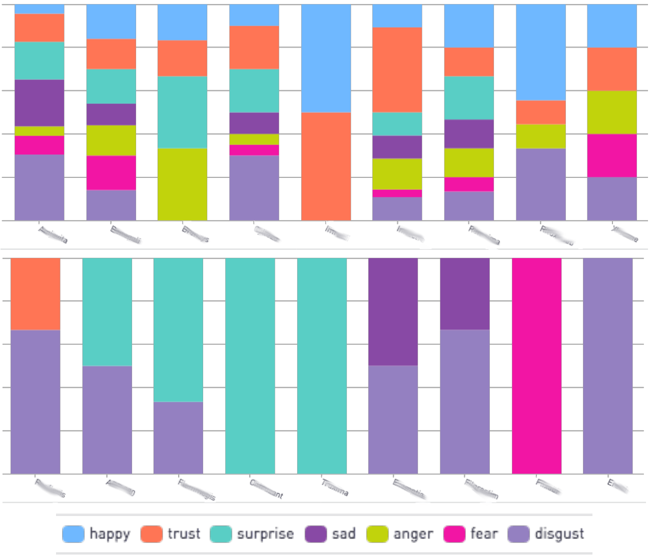
Joy

So grateful power has been restored & I was able to get my remicade infusion today. ? #HurricaneIrmaAftermath #autoimmune

Sadness

So far today: my insurance decided to stop covering my humira, my hip aching from the rain moving in, my foot is swollen, and I wanna cry.

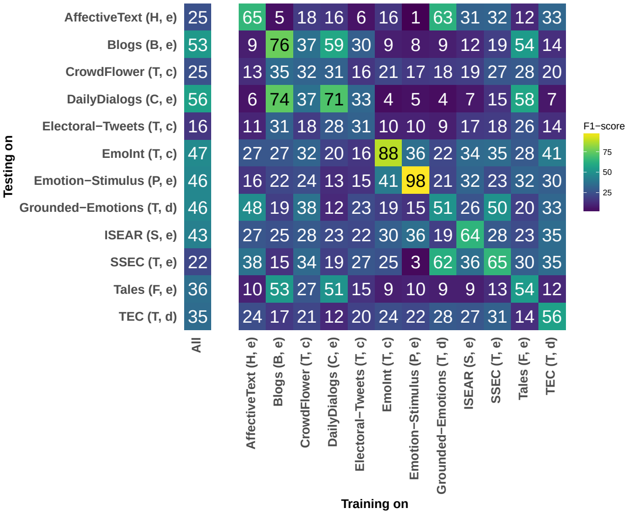
Comparison of Biosimilars



Challenges

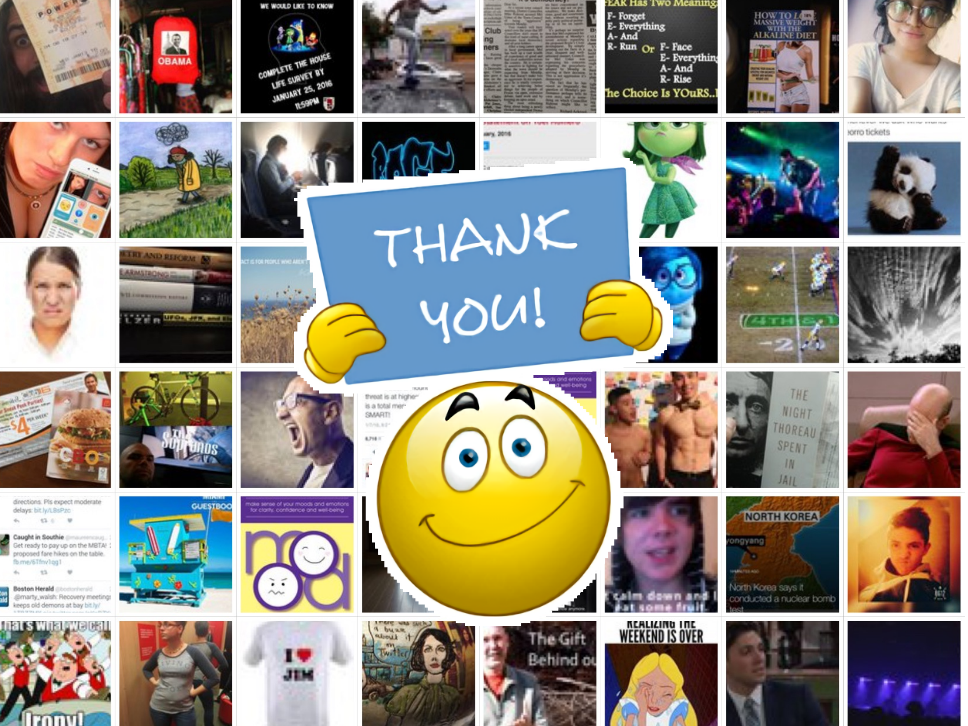
- Training on SSEC, testing on biosimilar data lead to $\approx 3\,\% F_1$
- Reannotation (of test, then of training data) is expensive and takes time (though that's what we do)
- Method selection for transfer learning from existing corpora is not solved

Transfer learning between corpora



Conclusion

- Emotion analysis backed well by psychology
- Transfer to natural language processing incomplete
- Models typically overfit to text source or domain
- Huge potential in applications across different domains


THANK YOU!
OBAMA
WE WOULD LIKE TO KNOW
COMPLETE THE HOUSE LIFE SURVEY BY JANUARY 25, 2016 11:59PM
FEAR Has Two Meanings
F- Forget E- Everything A- And R- Run Or F- Face E- Everything A- And R- Rise
The Choice Is YOuRS..!
HOW TO LOSE MASSIVE WEIGHT WITH THE ALKALINE DIET
THE NIGHT THOREAU SPENT IN JAIL
NORTH KOREA
North Korea says it conducted a nuclear bomb test
Caught in Southie
Get ready to pay up on the MBTA! proposed fare hikes on the table.
Boston Herald
keeps old demons at bay
GUESTBOOK
make sense of your moods and emotions for clarity, confidence and well-being
mood
I ♥ JIM
The Gift Behind
REALIZING THE WEEKEND IS OVER
Irony!

Special thanks to:

Laura Bostan

Evgeny Kim

Jeremy Barnes

More thanks to:

Sebastian Padó, Sabine Schulte im Walde, Hendrik Schuff, Saif Mohammad, Alexandra Balahur, Orphée De Clercq, Florian Barth, Sandra Murr, Matthias Hartung,...

DFG Deutsche Forschungsgemeinschaft

CRETA CENTER FOR REFLECTED TEXT ANALYTICS

Bundesministerium für Bildung und Forschung

SEMALYTIX