

Crowd Sourcing for Sentiment Analysis

ICA 2017

<http://vanatteveldt.com/ica2017-sentiment>

Is this positive or negative?

Cameron urges Greek euro exit plans

Is this positive or negative?

Cameron urges Greek euro exit plans

- Cameron is positive about the "plans"
- So is he positive about a Greek Euro exit?
- So is he negative about the Greek Euro?
- Is he maybe negative about Greece?
- And is he negative about the Euro?
- And what does Greece think of all this?

What is sentiment analysis?

- Sentiment: evaluative charge of subjective language
- Subjective language: reveals the inner state of author, persons in text

- Common approaches:
 - Human coding
 - Dictionary analysis
 - Machine learning

- Applications:
 - Analyzing reviews
 - Automatic stock market trading
 - Early warning systems / flagging potential negative posts
 - ...
- E.g. Pang/Lee 2008; Liu 2012; Wiebe 2004

Why is sentiment analysis difficult?

- "Terrorist Attacks: 250 Innocent Massacred By ISIS"
- "Saddam Hussein was executed by hanging"
- "This car has better mpg than my old Volvo"
- "Preacher who applauded Orlando mass killing asked to relocate"
- "Janeane Garafalo also was an interesting character ."
- "Brexit means brexit"
- "You should read the book!"

Why is sentiment analysis difficult?

- Evaluations are inherently subjective
- Evaluative language is creative and context-sensitive
 - (even more so than factual language)
- Evaluation inherently relational, but most data/tools undirected
 - Source *likes/dislikes* target

Task/domain specific sentiment analysis

- Sentiment value of words is context dependent
- Choice of words is context dependent
- Goal is often task dependent

Crowd sourcing 101

- Design (simple) task and instructions
- Upload units to e.g. crowdflower
- Random people code your units for e.g. \$.02 per unit
- Can use 'gold questions' to filter coders

Why would the crowd be good at sentiment?

- Task is "simple"
- We are (mostly) looking for a "gut feeling" response
- Subjective nature favours multiple codings
 - Spread as well as point estimate

(e.g. Benoit et al 2016; Haselmayer/Jenny 2016; Socher et al 2013)

Our approach

- ① Formal codebook for expert coding [hard!]
- ② Create gold standard for evaluation
- ③ Evaluate crowd coding on small sample
- ④ Run crowd coding on large corpus
- ⑤ Build and evaluate dictionary and ML model

Our approach

- ① Formal codebook for expert coding [hard!]
- ② Create gold standard for evaluation [hard!]
- ③ Evaluate crowd coding on small sample
- ④ Run crowd coding on large corpus
- ⑤ Build and evaluate dictionary and ML model

Our approach

- ① Formal codebook for expert coding [hard!]
- ② Create gold standard for evaluation [hard!]
- ③ Evaluate crowd coding on small sample [hard!]
- ④ Run crowd coding on large corpus
- ⑤ Build and evaluate dictionary and ML model

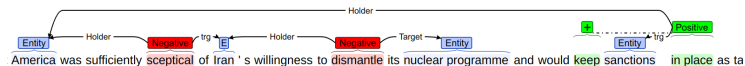
Our approach

- ① Formal codebook for expert coding [hard!]
- ② Create gold standard for evaluation [hard!]
- ③ Evaluate crowd coding on small sample [hard!]
- ④ ~~Run crowd coding on large corpus~~
- ⑤ ~~Build and evaluate dictionary and ML model~~

So what was the problem?

- We want to code relational sentiment
 - target: Sentiment **about** someone
 - holder: Sentiment **from** someone
- What are the nodes?
- How to explain difference between sentiment, target, and holder to the crowd?

Aside: creating the gold standard



Attempt 1: let crowd pick target word(s)

(1) Carefully read the following sentence:

We do not want to intimidate others , nor force our ideas on them .

(2) Please **indicate who or what the opinion is about** by highlighting these words in the box below.

In a sentence like *"David Cameron defended his announcement"*, Cameron has an opinion about "his announcement", so high **Don't highlight the opinion itself** and **don't highlight the person who expresses the opinion**, so do not highlight the words "c" example.

We do not want to intimidate others , nor force our ideas on them .

Attempt 1: Problems

	n	holder	sentiment	target
1	271	41	50	33
2	159	34	41	21
3	86	14	24	14
4	57	8	24	20
5	53	3	12	13
6	34	0	10	18
7	26	0	9	13
8	11	1	2	10
9	23	0	5	9
10	4	0	5	7

- Coders could not distinguish target, holder, sentiment

Attempt 2: select target from NPs

- Identify all NPs in the sentence
- Let crowd coder select from NP:

Please read the following sentence about the longer-term future:

Then there is the question of the longer-term future for Europe .

Is someone (or the author) positive or negative about the longer-term future in this sentence?

- No, no one is no positive or negative about the longer-term future
- Yes, someone is positive about the longer-term future
- Yes, someone is negative about the longer-term future

Attempt 2: problems

- Most NPs are not potential targets
- Overall kappa .62
- Agreement mostly on what wasn't a target
- First filter pass to throw out 'junk' didn't yield enough

Attempt 3: select holder from NEs

- Can use named entity recognition to 'filter' NPs
 - Agents, organizations, locations (=countries)

Then there is the question of the longer-term future for Europe .

The country is outside the euro zone but about 40 percent of its exports go to the single currency bloc .

Two of the largest foreign investors in India are British - Vodafone and BP .

Among ministers , there was fury at the way Labour had , in their view , switched position to outflank the Government on its most vul

They speak about this shift in Brussels and Paris and certainly in Athens , Rome and Madrid .

President François Hollande also said that France could be willing to supply arms at some point .

Mr Hollande 's first official foreign trip will be to Berlin for a meeting with Ms Merkel after his inauguration on May 15 .

I am not a British isolationist but I do want a better deal for Britain , " he said .

Regional and Western governments have expressed growing concern about the security threat from extremists and organised crime .

- Recall of gold standard holders/targets:

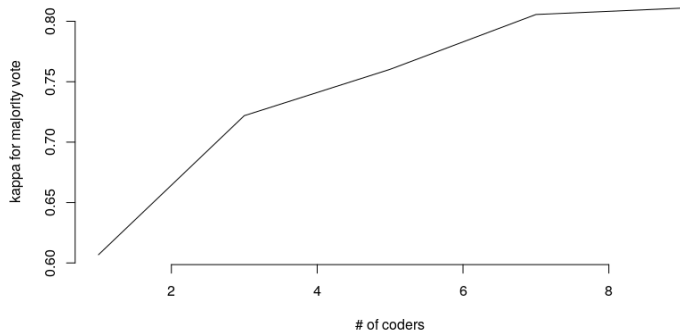
type	holder	sentiment	target
NER	0.67	0.01	0.51
NP	0.96	0.45	0.88

Please read the following sentence about **India**:

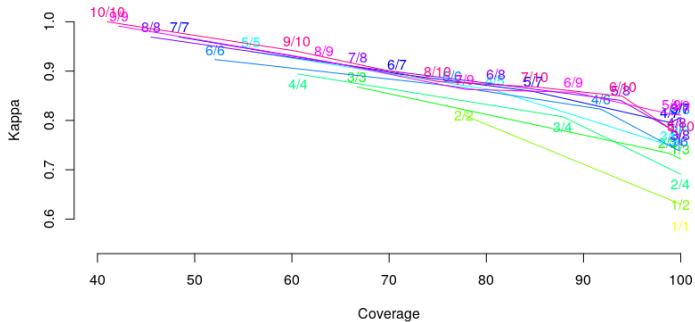
Two of the largest foreign investors in **India** are British - Vodafone and BP .

Is India positive or negative about something or someone in this sentence?

- No, **India** is not positive or negative about something or someone in this sentence
- Yes, **India** is positive about something or someone in this sentence
- Yes, **India** is negative about something or someone in this sentence



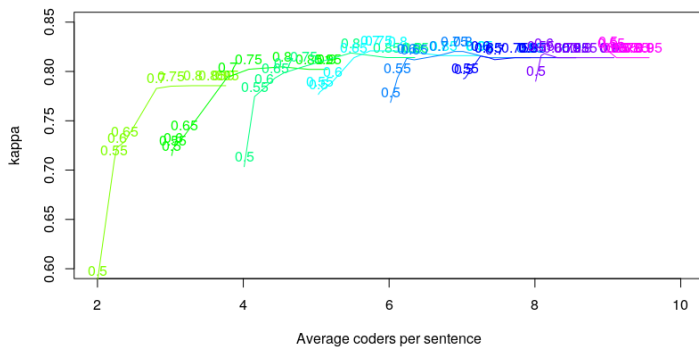
Coverage/kappa for holder-sentiment



Dynamic judgments

- Some sentences are harder than others
- Crowdfunder can automatically add covers if confidence is low
- Can we use this to save money?

Performance with dynamic judgments



Second step: select target from NPs

Please read the following sentence in which **France** expresses a **positive** opinion:

[President François Hollande](#) also said that **France** could be willing to supply [arms](#) at [some point](#).

About who or what is France positive?

- President François Hollande
- arms
- some point
- NONE: France is not positive about any of these

Second step: select target from NPs

- Only evaluated on 26 sentences for now
- 20 were correct, but most other reasonable
 - Problems often non-NP target ("to supply arms")
- Need to test on larger scale

Conclusion

- Sentiment is (still) hard :)
- Crowd sourcing is (still) a promising technique
- Lessons:
 - Task needs to be simple multiple choice
 - Gold questions are very important
 - Work needs to be done on generating potential holders/targets
- Plans:
 - Run larger scale evaluation
 - Crowd code larger corpus
 - Experiment with dictionary / ML