cross domain coutnerargument retrieval

Bachelor thesis of Janko Götze

Motivation

- Strengthening arguments by preparing for possible counters
- Also, the opposite: preparing against an argument by gathering counters
- Unbiased help in decision-making
- Possible applications in Marketing, Education, Research and others

Research Questions

- Wachsmuth et al (2018) defined a counter-argument as the one that "invokes the same aspects as the (input) argument while having the opposite stance". Their approach to find the best counterargument is to select the most semantically similar to the given argument and simultaneously most dissimilar to it in terms of stance.

In the context of this idea, I pose the following research questions:

1. To what extent can we effectively predict persuasive counter-arguments in online discussions using the proposed approach?

2. Can this approach be adapted or modified to improve its effectiveness in online persuasive discussions?

Main Inspiration

- "invokes the same aspects as the (input) argument while having the opposite stance"
- Concrete: similar words, but different stance
- Realized through word similarity and embedding similarity
- Used to form a score

Change my view (CMV)

- Subreddit (online forum)

- Someone posts a claim

- Others try to convince them otherwise

- If convinced comments can be awarded with deltas

Example

Datasets

ArguArna Corpus	Cange my view
idebate.net	reddit.com
Topics get reviewed	Everyone can post
Clear structure	General rules
One argument and one counterargument	Many comments with potentially many delta comments
Generally objective	Highly subjective

Experiments

- Recreation of Wachsmuth method in python
  - Originally in Java with proprietary tools
  - Not one to one replica achieved
  - About 10% less accuracy
  - Probably because of tokenizers and embeddings

- Change my view
  - Only threads with delta comments
  - Two settings:
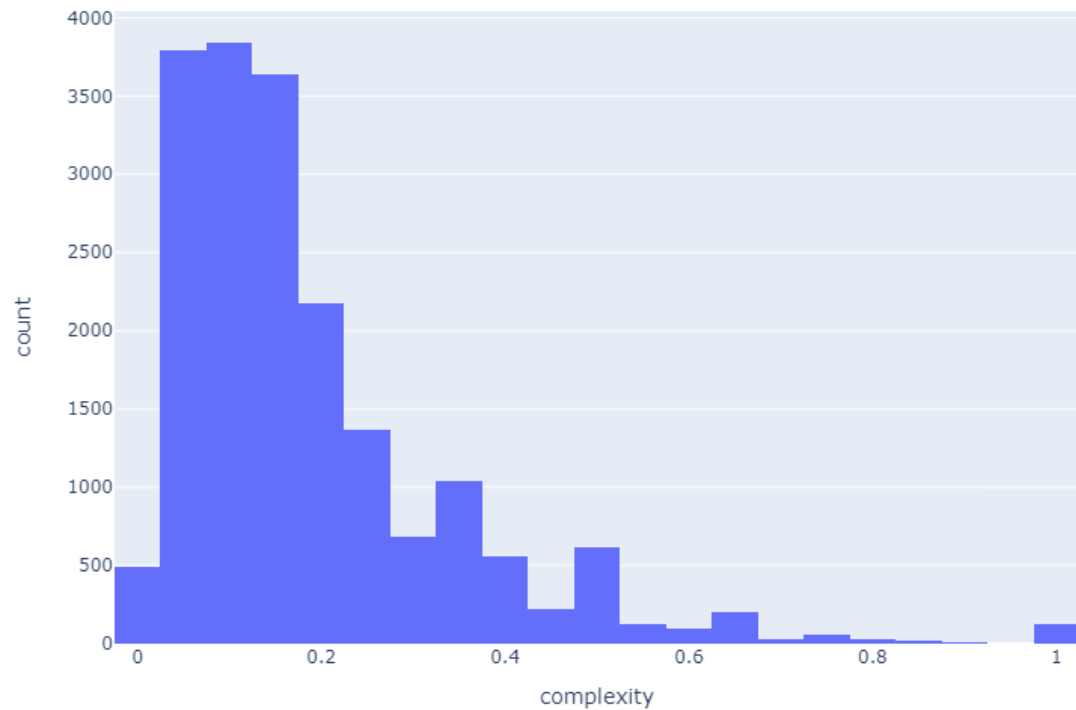   - Only root comments
    - All comments

Results

- Pretty bad performance
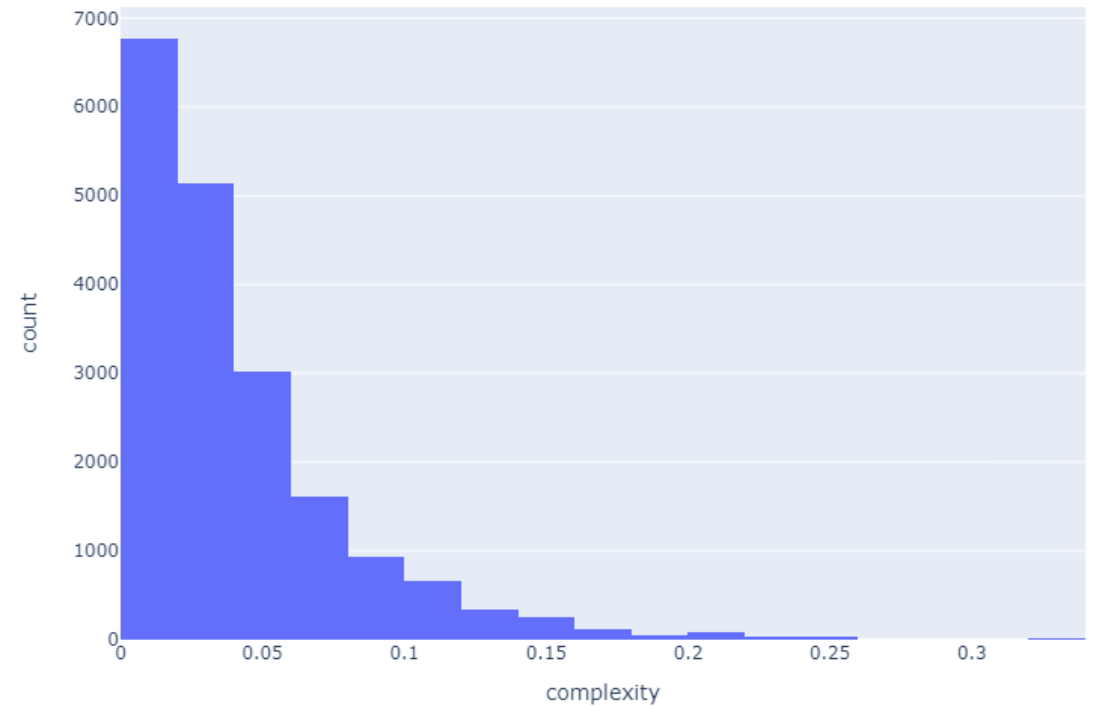- Better than random
- Beat by baseline heuristics

Accuracy	Root comments	All comments
My Method	23.85%	7.83%
Random	18.20%	4.10%
Longest	32.53%	20.35%
Earliest	25.85%	8.93%

Dataset distribution

Complexity = delta comments / total comments
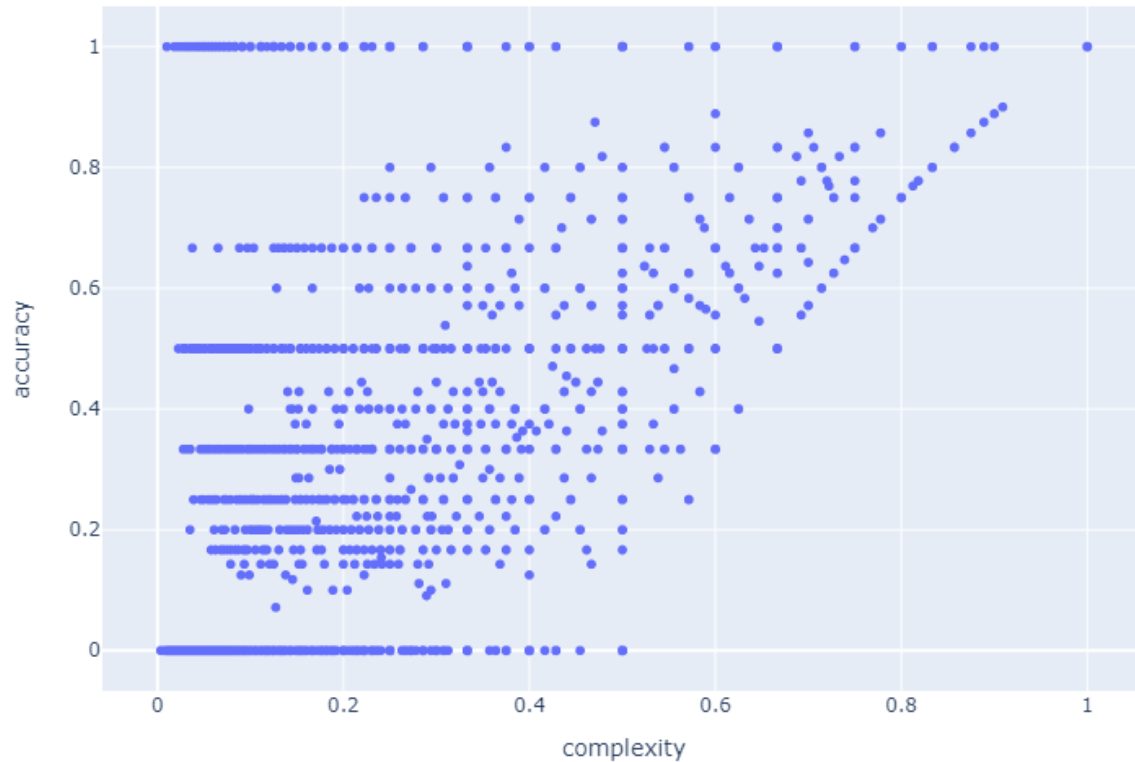

Only Root comments

All comments

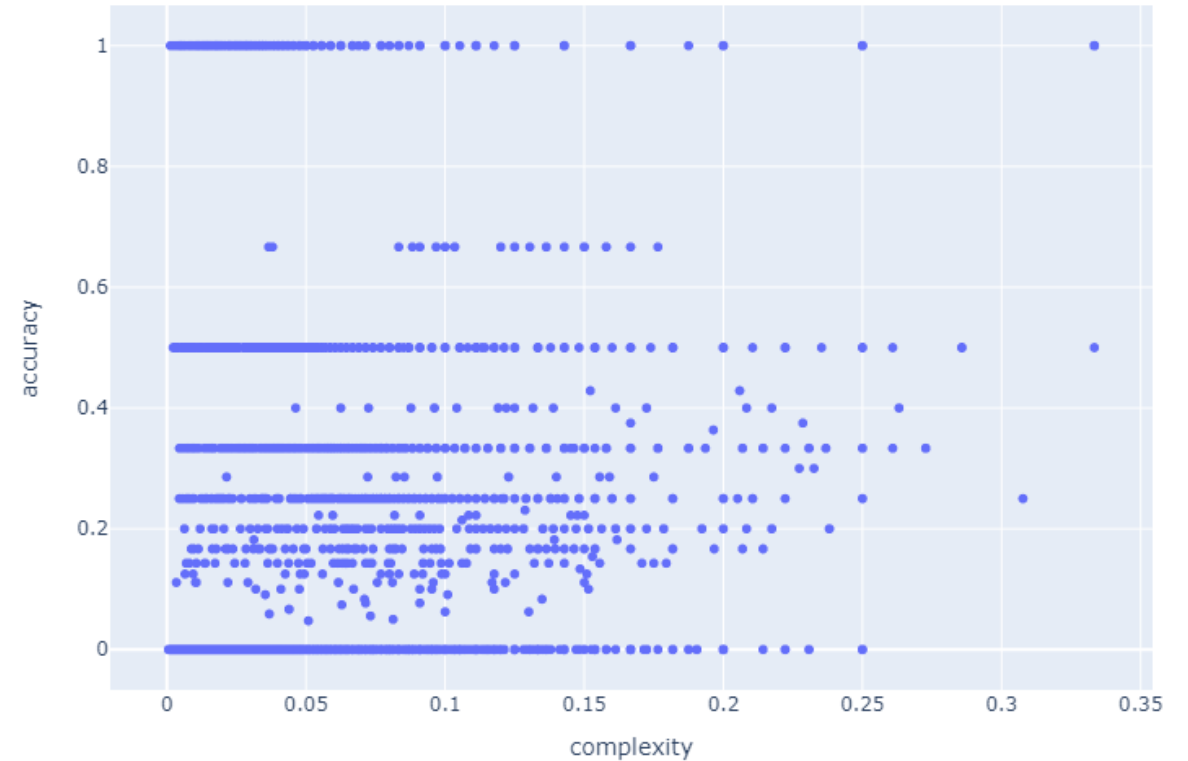# Evaluation

Clear groups at low fractions: 0.25, 0.33, 0.5

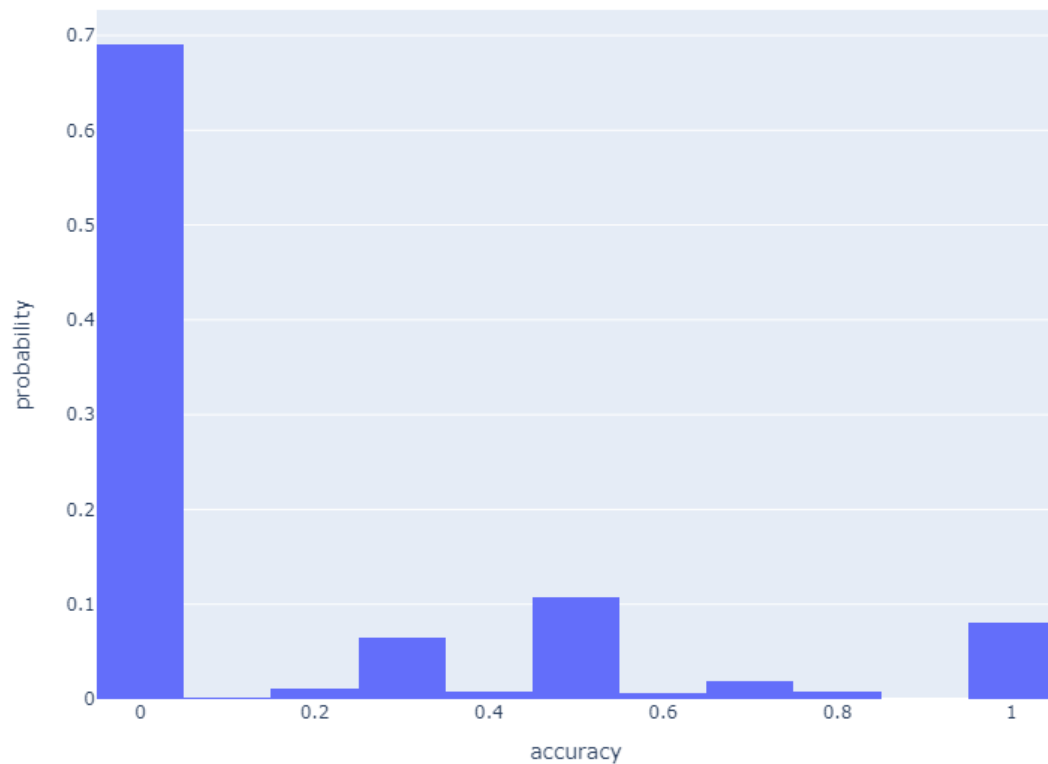Might seem reasonable, until we look at the distribution
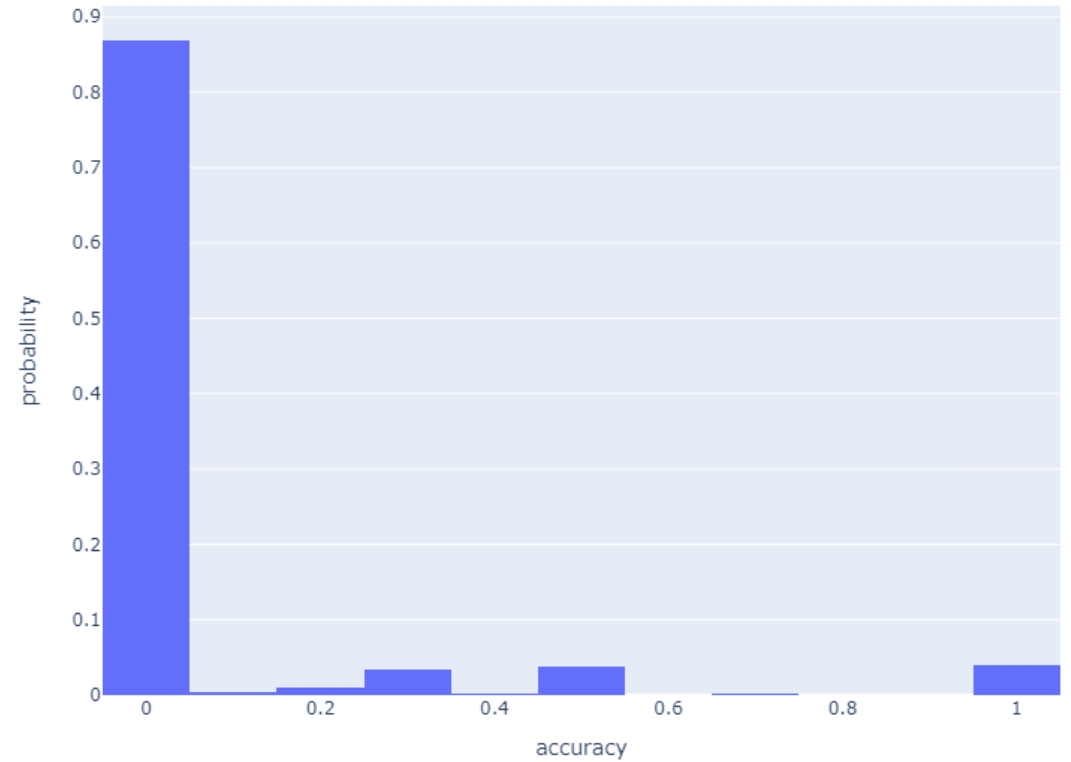

Only Root comments

All comments

Evaluation

For roots: nearly 70% have 0% accuracy

For all comments: over 85% have 0% accuracy

Only Root comments

All comments

Conclusion

- Did not work effectively for online discussions
- Various potential factors
  - Length of Thread
  - Length of comments
  - Complexity
  - Other nuances like citations of OP

- More Investigation needed

TODOs / future work

In depth error analysis

Improve example evaluation

Try different inprovments like prefiltering and multiple passes