Grimjack at Touché 2022

Advanced IR, Winter Semester 2021/22
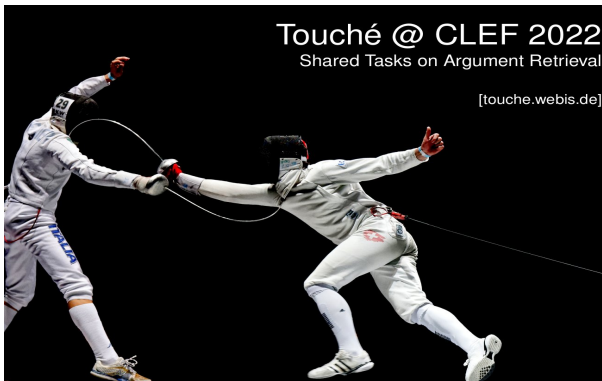
Johannes Huck   Jan Heinrich Reimer

Martin Luther University Halle-Wittenberg

January 29, 2022

Task at hand

- ▶ Task 2 of Touché: Argument Retrieval
- ▶ Argument Retrieval for Comparative Questions
- ▶ Task: Retrieve relevant passages to answer comparative questions and detect their stance w.r.t the objects
- ▶ Data: > 1 million text passages

General approach

- ▶ Programmed in Python
  - ▶ Easy to use
  - ▶ High readability
  - ▶ Many IR libabries avaiable
- ▶ Three modules: Search, Search-all and Evaluate-all
- ▶ Pipeline consists of
  - ▶ Query expander
  - ▶ Initial Retrieval
  - ▶ Argument quality and stance tagging
  - ▶ Reranking
- ▶ Each pipeline step can be activated/deactivated

Pipeline

Query Expander

- Expanding queries with synonyms of comparative objects
- Two Different approaches
    - Based on embeddings with glove
    - Based on language model T0 [San+21]
    - We ask "What are synonyms of the word <token> ?"
- With T0 also new queries from narrative and description
- We ask "Extract a query: <text>"
- Combining all new queries with OR
- Retrieving ranked list of passages with this new query

Argument quality tagging

- Extracting arguments with TARGER [Che+19]
- For each argument we want to know the quality w.r.t. the topic
- We use the IBM Debater API [Tol+19] to get the argument quality
- Returns a score between 0 and 1
- 0 means lowest quality and 1 highest quality

Example

Arg: Cars should only provide assisted driving, not complete autonomy
Topic: We should further explore the development of autonomous vehicles
Score: 0.7256

Argument stance tagging

- Next we want to know the stance w.r.t. the topic
- Two different approaches
  - Based on IBM Debater API [Bar+17]
  - Based on T0
  - We ask "Is this sentence positive or negative with respect to <comparative_object>? <sentence>"
- It is also possible to expand with sentiments
- Calculating the multi target stance
  - Calculate the difference between objects
  - Use a threshold
  - Convert T0s output to numbers

Reranking

- Foo

Final Remarks

- Approach is very flexible
- We investigate influence of components w.r.t the retrieval score
- Stance classification may be better with Roberta approach
- We cannot distinguish between neutral and no stance
- We investigate how reranking influences the retrieval score
- T0 solves a lot of IR Task
- Is it possible to only use T0 for retrieval?

Thank you!

References

Bar-Haim, Roy et al. (2017). "Stance Classification of Context-Dependent Claims". In: *EACL*.

Chernodub, Artem N. et al. (2019). "TARGER: Neural Argument Mining at Your Fingertips". In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28 - August 2, 2019, Volume 3: System Demonstrations*. Ed. by Marta R. Costa-jussà et al. Association for Computational Linguistics, pp. 195–200.

Sanh, Victor et al. (2021). "Multitask Prompted Training Enables Zero-Shot Task Generalization". In: *CoRR* abs/2110.08207. arXiv: `2110.08207`.

Toledo, Assaf et al. (2019). "Automatic Argument Quality Assessment - New Datasets and Methods". In: *EMNLP*.