TARGER: Neural Argument Mining at Your Fingertips

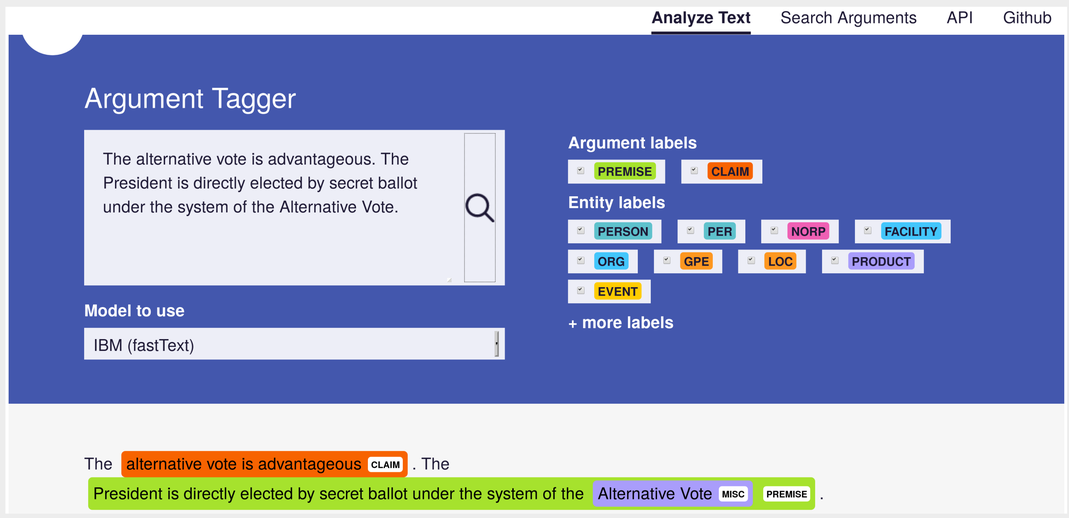
Artem Chernodub[1] Oleksiy Oliynyk[2] Philipp Heidenreich[2]
Alexander Bondarenko[3] Matthias Hagen[3] Chris Biemann[2] Alexander Panchenko[2]

[1]artem.chernodub@grammarly.com [2]first.last@informatik.uni-hamburg.de [3]first.last@informatik.uni-halle.de

Abstract

We present TARGER, an open source neural argument mining framework, which comes with a web interface and API for neural argument mining and retrieval. It can tag arguments in free texts and retrieve arguments from the Common Crawl-based DepCC corpus (Panchenko et al., 2018).

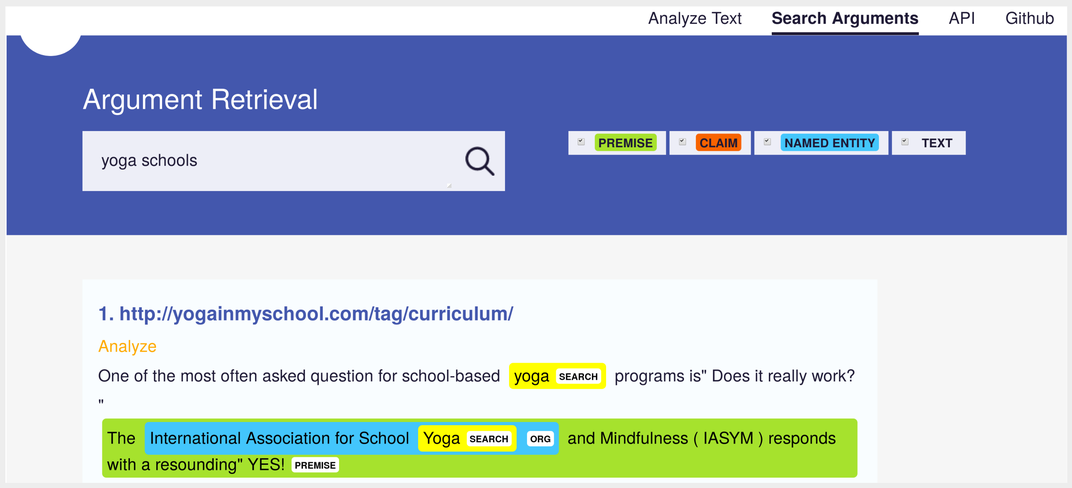
Analyze Input Text



TARGER's Assets

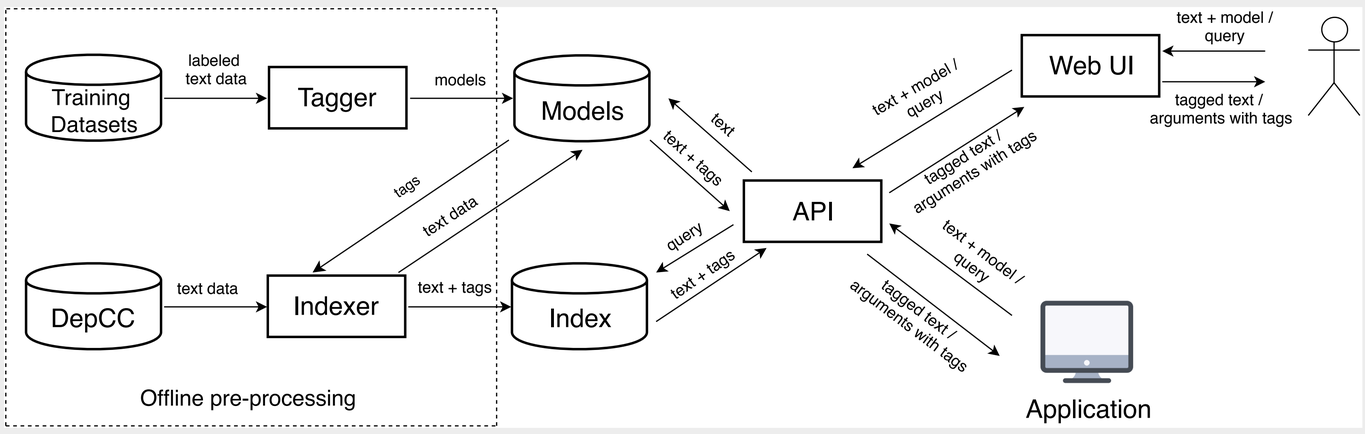
- Easily accessible web interface and RESTful API
- Simple to incorporate into existing NLP pipelines
- Web-scale retrieval of argumentative sentences
- Various models pretrained on diverse datasets

Search Common Crawl for Arguments



http://ltdemos.informatik.uni-hamburg.de/argsearch

Modular Architecture



- BiLSTM-CNN-CRF neural tagger for argument components identification (Ma and Hovy, 2016)
- GloVe (Pennington et al., 2014), fastText (Mikolov et al., 2018), and dependency-based embeddings (Levy and Goldberg, 2014)
- Pre-trained on persuassive essays (Eger et al., 2017), web discourse (Habernal and Gurevych, 2017), and IBM Debater data (Levy et al., 2018)
- Implemented in Python 3.6 / PyTorch 1.0

Conclusions

- User-friendly web interface and easily accessible RESTful API for argument mining
- Easily re-trainable models on standard datasets
- Next steps:
  - Integrating models based on ELMo (Peters et al., 2018) and BERT (Devlin et al., 2019)
  - Tackling the problem of domain-specific argument tagging

References

Devlin, J., Chang, M., Lee, K., and Toutanova, K. (2019). BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019*.
Eger, S., Daxenberger, J., and Gurevych, I. (2017). Neural End-to-End Learning for Computational Argumentation Mining. In *Proceedings of ACL 2017*.
Habernal, I. and Gurevych, I. (2017). Argumentation Mining in User-Generated Web Discourse. *Computational Linguistics*, 43(1).
Levy, O. and Goldberg, Y. (2014). Dependency-Based Word Embeddings. In *Proceedings of ACL 2014*.
Levy, R., Bogin, B., Gretz, S., Aharonov, R., and Slonim, N. (2018). Towards an Argumentative Content Search Engine Using Weak Supervision. In *Proceedings of COLING 2018*.
Ma, X. and Hovy, E. H. (2016). End-to-End Sequence Labeling via Bi-directional LSTM-CNNs-CRF. In *Proceedings of ACL 2016*.
Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., and Joulin, A. (2018). Advances in Pre-Training Distributed Word Representations. In *Proceedings of LREC 2018*.
Panchenko, A., Ruppert, E., Faralli, S., Ponzetto, S. P., and Biemann, C. (2018). Building a Web-Scale Dependency-Parsed Corpus from CommonCrawl. In *Proceedings of LREC 2018*.
Pennington, J., Socher, R., and Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In *Proceedings of EMNLP 2014*.
Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., and Zettlemoyer, L. (2018). Deep Contextualized Word Representations. In *Proceedings of NAACL-HLT 2018*.

This work was partially supported by the German Academic Exchange Service (DAAD) through the short-term research grant 57314022 "Argument Mining in Web Documents using Orthogonal Simple Recurrent Networks" and by the Deutsche Forschungsgemeinschaft (DFG) through the project "ACQuA: Answering Comparative Questions with Arguments" (grants BI 1544/7-1 and HA 5851/2-1) as part of the priority program "RATIO: Robust Argumentation Machines" (SPP 1999).