Argumentation Mining

Filip Boltuzic

2018
March

1 Definiton of Argumentation Mining

Argument(ation) Mining is the automatic identification of the argumentative structure contained withing in a piece of language [Lawrence and Reed, 2017].

Argument mining is automatic extraction of arguments from natural text [Aker et al., 2017].

2 Tasks of Argumentation Mining

1. Clustering recurring arguments [Boltužić and Šnajder, 2015, Misra et al., 2017]
2. Recognizing argument schemes [Feng and Hirst, 2011].
3. Prediction of structure (connecting premises to claims) [Aker et al., 2017, Lawrence and Reed, 2017]

2.1 Identifying argumentative segments in text

Unit segmentation consists in the splitting of a text into its argumentative segments (ADU) and their non-argumentative counterparts.

[Ajjour et al., 2017]

[Persing and Ng, 2016] rely on handcrafted features based on the parse tree of a sentence to identify segments.

[Stab, 2017] uses sequence modeling and sophisticated features to classify the argumentativeness of each single word based on its surrounding words.

[Eger et al., 2017] employ a deep learning architecture using different features based on the entire essagt

[Al Khatib et al., 2016] have a rule-based where they suggest where the arguments should be split before the actual argument annotation (annotators could merge arguments back).

[Aker et al., 2017] determine if a sentence is a claim, premise or none. They work on a sentence boundary.

2.2 Prediction of structure

[Lawrence and Reed, 2017] have annotated debates on "Moral Maze" and created argument diagrams via AIFDB. They aim to recognize the support relation from text (inference or non-inference).

[Aker et al., 2017] use claim-premise pairs and go full Carteisan on them, making negative examples for those who aren't linked in the gold set. They work on the [Stab and Gurevych, 2017, Aharoni et al., 2014] datasets.

3 Unsupervised approaches to Argumentation Mining

Lack of large datasets for argumentation mining is one of the largest concerns of the community.

[Habernal and Gurevych, 2015] try to use unsupervised features for better argument component identification from online debate portals.

[Al-Khatib et al., 2016] apply distant supervision to automatically create a large annotated corpus from online debate portals with argumentative and non-argumentative segments from several domains.

[Lawrence and Reed, 2017] try to use web search in combination with *therefore and because* discourse indicators in addition to some other filtering. They make their own premise-conclusion pairs by searching the web for the discourse marker and then use LDA to predict support/non-support relations.

4 Predicting support relations

Predicting support relations is similar to textual entailment, but involves more contextual knowledge and common-sense reasoning since the semantic distance is greater. Also, it is not strictly a logical relation and (with a well-defined hypothesis–text relation), but (usually) there is a direction defined.

[Lawrence and Reed, 2017] constructs a corpus using web-search and a gold set then does supervised classification whether a sentence supports (infers) another.

5 Cross-domain argumentation mining

[Ajjour et al., 2017] do argumentative unit segmentation on three corpuses: Habernal's Web Discourse, Stab's Essay corpus, and Editorials to show how cross-domain argumentative unit segmentation is a huge problem as it is defined today and even end with open questions about how should segmentation and argumentative units be defined.

References

[Aharoni et al., 2014] Aharoni, E., Polnarov, A., Lavee, T., Hershcovich, D., Levy, R., Rinott, R., Gutfreund, D., and Slonim, N. (2014). A benchmark dataset for automatic detection of claims and evidence in the context of controversial topics. In *Proceedings of the First Workshop on Argumentation Mining*, pages 64–68.

[Ajjour et al., 2017] Ajjour, Y., Chen, W.-F., Kiesel, J., Wachsmuth, H., and Stein, B. (2017). Unit Segmentation of Argumentative Texts. pages 118–128.

[Aker et al., 2017] Aker, A., Sliwa, A., Ma, Y., Liu, R., Borad, N., Ziyaei, S. F., and Ghbadi, M. (2017). What works and what does not: Classifier and feature analysis for argument mining. pages 91–96.

[Al-Khatib et al., 2016] Al-Khatib, K., Wachsmuth, H., Hagen, M., Köhler, J., and Stein, B. (2016). Cross-domain mining of argumentative text through distant supervision. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1395–1404.

[Al Khatib et al., 2016] Al Khatib, K., Wachsmuth, H., Kiesel, J., Hagen, M., and Stein, B. (2016). A news editorial corpus for mining argumentation strategies. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 3433–3443.

[Boltužić and Šnajder, 2015] Boltužić, F. and Šnajder, J. (2015). Identifying prominent arguments in online debates using semantic textual similarity. In *Proceedings of the 2nd Workshop on Argumentation Mining*, pages 110–115.

[Eger et al., 2017] Eger, S., Daxenberger, J., and Gurevych, I. (2017). Neural end-to-end learning for computational argumentation mining. *arXiv preprint arXiv:1704.06104*.

[Feng and Hirst, 2011] Feng, V. W. and Hirst, G. (2011). Classifying arguments by scheme. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 987–996. Association for Computational Linguistics.

[Habernal and Gurevych, 2015] Habernal, I. and Gurevych, I. (2015). Exploiting debate portals for semi-supervised argumentation mining in user-generated web discourse. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2127–2137.

[Lawrence and Reed, 2017] Lawrence, J. and Reed, C. (2017). Mining Argumentative Structure from Natural Language text using Automatically Generated Premise-Conclusion Topic Models. pages 39–48.

[Misra et al., 2017] Misra, A., Anand, P., Tree, J. E. F., and Walker, M. (2017). Using summarization to discover argument facets in online ideological dialog. *arXiv preprint arXiv:1709.00662*.

[Persing and Ng, 2016] Persing, I. and Ng, V. (2016). End-to-end argumentation mining in student essays. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1384–1394.

[Stab and Gurevych, 2017] Stab, C. and Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.

[Stab, 2017] Stab, C. M. E. (2017). *Argumentative writing support by means of natural language processing*. PhD thesis, Technische Universität Darmstadt.