

Section-wise Stance Detection: Analyzing Article Structure and Length Asymmetries

Ali Mhrez^a, Tibor Tajti^b

^aUniversity of Debrecen, Doctoral School of Informatics
mhrez.ali@inf.unideb.hu

^bUniversity of Debrecen, Faculty of Informatics and Eszterházy Károly Catholic University
tajti.tibor@inf.unideb.hu, tajti.tibor@uni-eszterhazy.hu

Abstract

The rapid, unmoderated exchange of news on social media platforms significantly facilitates the spread of inaccurate information, notably fake news. Stance detection, the task of classifying an article’s stance toward a headline, is a crucial step in combating this phenomenon [6]. Existing machine learning models, including Transformer-based architectures, face two primary challenges in fake news stance detection:

1. **Article Length:** News articles can span thousands of words, often containing irrelevant or varied information.
2. **Length Asymmetry:** The inherent imbalance between short headlines and long articles may bias a model’s attention, making it harder for the model to learn the underlying relationship between them.

This paper investigates these length challenges and asymmetries to identify optimal processing techniques. Specifically, we explore (1) the utility and potential detriment of different article text sections for accurate stance detection, and (2) which article sections (lead, body, or conclusion) contain primary stance cues. To our knowledge, this is the first systematic and empirical study to investigate how stance-related cues are distributed across an article’s core structural sections.

The researchers evaluated two Transformer-based models: ArBERTv2 [1] for Arabic and RoBERTa [5] for English. These were tested across three benchmark datasets: AraStance (4,063 instances) [2], UnifiedFC (3,042 instances) [3], and FNC-1 (75,385 instances) [4]. The study employed several experimental configurations to probe article utility:

- **Lead-N:** Combining the headline with the first N tokens of the article.
- **Conclusion-N:** Combining the headline with the final N tokens of the article.
- **Sliding Window Segments:** Combining the headline with a specific segment of the article.
- **Noheadline-N:** An ablation study using only the article text to quantify the headline’s contribution.

Performance was primarily measured using the Macro F1-score to account for significant class imbalances within the datasets.

Our comprehensive experiments reveal several primary insights regarding article structure and model performance:

1. **The Headline is Paramount:** Across all datasets, including the headline consistently and dramatically improves performance. In the FNC-1 dataset, the Macro F1-score dropped from 78.63% (Lead-512) to just 28.16% (Noheadline-512), proving that articles cannot be accurately classified without the headline’s context.
2. **The Inverted Pyramid Style:** Stance cues are most concentrated in the article lead. Models using only the lead sections achieved peak Macro F1 scores of 77.58% (AraStance), 47.35% (UnifiedFC), and 78.63% (FNC-1).
3. **The Importance of Conclusions:** While the lead is generally superior, core arguments often appear at the end of articles. Performance in Conclusion-N settings improved consistently as the token window increased, suggesting that relevant information at the end of a document is more distributed.
4. **Optimal Input Length:** The optimal input length is not universal. For example, AraStance peaked at 256 tokens, while FNC-1 and UnifiedFC benefited from the maximum 512-token window.

This study provides an empirical foundation for advancing stance detection by proving that simplistic truncation (Lead-only) is insufficient and potentially detrimental, as it ignores vital cues located in article conclusions. Our results highlight the need for dynamic text processing strategies, such as segment aggregation or summarization, that can bridge the length gap between headlines and articles while capturing information dispersed throughout the text. To facilitate reproducibility and encourage further research, we have made our complete codebase and dataset splits publicly available.

References

- [1] M. ABDUL-MAGEED, A. ELMADANY, E. M. B. NAGOUDI: *ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic*, in: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Online: Association for Computational Linguistics, Aug. 2021, pp. 7088–7105, DOI: [10.18653/v1/2021.acl-long.551](https://doi.org/10.18653/v1/2021.acl-long.551).
- [2] T. ALHINDI, A. ALABDULKARIM, A. ALSHEHRI, M. ABDUL-MAGEED, P. NAKOV: *AraStance: A Multi-Country and Multi-Domain Dataset of Arabic Stance Detection for Fact Checking*, in: Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda, Online: Association for Computational Linguistics, June 2021, pp. 57–65, DOI: [10.18653/v1/2021.nlp4if-1.9](https://doi.org/10.18653/v1/2021.nlp4if-1.9).
- [3] R. BALY, M. MOHTARAMI, J. GLASS, L. MÀRQUEZ, A. MOSCHITTI, P. NAKOV: *Integrating Stance Detection and Fact Checking in a Unified Corpus*, in: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), New Orleans, Louisiana: Association for Computational Linguistics, June 2018, pp. 21–27, DOI: [10.18653/v1/N18-2004](https://doi.org/10.18653/v1/N18-2004).
- [4] B. GALBRAITH, D. RAO: *Stance Detection dataset for FNC-1*, Accessed: 2026-01-20, 2017, URL: <https://github.com/FakeNewsChallenge/fnc-1>.
- [5] Y. LIU, M. OTT, N. GOYAL, J. DU, M. JOSHI, D. CHEN, O. LEVY, M. LEWIS, L. ZETTLEMOYER, V. STOYANOV: *RoBERTa: A Robustly Optimized BERT Pretraining Approach*, arXiv preprint (2019), arXiv: [1907.11692](https://arxiv.org/abs/1907.11692) [cs.CL].
- [6] D. POMERLEAU, D. RAO: *Fake news challenge*, Accessed: 2026-01-20, 2017, URL: <http://www.fakenewschallenge.org/>.