















- of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), Association for Computational Linguistics, Jun. 2018, pp. 2227–2237.
- [17] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018, doi: 10.48550/arXiv.1810.04805.
- [18] A. Vaswani *et al.*, “Attention is all you need,” *Adv Neural Inf Process Syst*, vol. 30, 2017.
- [19] S. Zhu, S. Li, and G. Zhou, “Adversarial attention modeling for multi-dimensional emotion regression,” in *Proceedings of the 57th annual meeting of the association for computational linguistics*, 2019, pp. 471–480. doi: 10.18653/v1/P19-1045.
- [20] Y. Yang, D. Zhou, Y. He, and M. Zhang, “Interpretable relevant emotion ranking with event-driven attention,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, 2019, pp. 177–187. doi: 10.18653/v1/D19-1017.
- [21] Y. Liu *et al.*, “Roberta: A robustly optimized bert pretraining approach,” *arXiv preprint arXiv:1907.11692*, 2019, doi: 10.48550/arXiv.1907.11692.
- [22] Z. Wu, X. Zhang, T. Zhi-Xuan, J. Zaki, and D. C. Ong, “Attending to emotional narratives,” in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, 2019, pp. 648–654. doi: 10.1109/ACII.2019.8925497.
- [23] D. Cortiz, “Exploring transformers in emotion recognition: a comparison of bert, distillbert, roberta, xlnet and electra,” *arXiv preprint arXiv:2104.02041*, 2021, doi: 10.48550/arXiv.2104.02041.
- [24] V. Francisco, P. Gervás, and F. Peinado, “Ontological reasoning to configure emotional voice synthesis,” in *Web Reasoning and Rule Systems: First International Conference, RR 2007, Innsbruck, Austria, June 7-8, 2007. Proceedings 1*, Springer, 2007, pp. 88–102. doi: 10.1007/978-3-540-72982-2\_7.
- [25] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *J Behav Ther Exp Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994, doi: 10.1016/0005-7916(94)90063-9.
- [26] James Carney and Cole Robertson, “4000 stories with sentiment analysis dataset.” Accessed: May 19, 2023. [Online]. Available: [https://brunel.figshare.com/articles/dataset/4000\\_stories\\_with\\_sentiment\\_analysis\\_dataset/7712540](https://brunel.figshare.com/articles/dataset/4000_stories_with_sentiment_analysis_dataset/7712540)
- [27] M.-T. Luong, H. Pham, and C. D. Manning, “Effective approaches to attention-based neural machine translation,” *arXiv preprint arXiv:1508.04025*, 2015, doi: 10.48550/arXiv.1508.04025.
- [28] X. Liu, P. He, W. Chen, and J. Gao, “Multi-task deep neural networks for natural language understanding,” *arXiv preprint arXiv:1901.11504*, 2019, doi: 10.48550/arXiv.1901.11504.
- [29] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, “Albert: A lite bert for self-supervised learning of language representations,” *arXiv preprint arXiv:1909.11942*, 2019, doi: 10.48550/arXiv.1909.11942.
- [30] J. Wei and K. Zou, “Eda: Easy data augmentation techniques for boosting performance on text classification tasks,” *arXiv preprint arXiv:1901.11196*, 2019, doi: 10.48550/arXiv.1901.11196.