

REFERENCE

Ba, J.L., Kiros, J.R., Hinton, G.E., 2016. *Layer Normalization*. arXiv:1607.06450[cs,stat].

Cai, Z., Tu, L., Gimpel, K., 2017. *Pay Attention to the Ending: Strong Neural Baselines for the ROC Story Cloze Task*, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Presented at the Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Association for Computational Linguistics, Vancouver, Canada, pp. 616–622. <https://doi.org/10.18653/v1/P17-2097>

Chaturvedi, S., Peng, H., Roth, D., 2017. *Story Comprehension for Predicting What Happens Next*, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Presented at the Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Copenhagen, Denmark, pp. 1603–1614. <https://doi.org/10.18653/v1/D17-1168>

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv:1810.04805 [cs].

Ethayarajh, K., 2019. *How Contextual are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings*, 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). Presented at the 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, Hong Kong, China, pp. 55–65. <https://doi.org/10.18653/v1/D19-1006>

Li, Y., Xu, L., Tian, F., Jiang, L., Zhong, X., Chen, E., n.d. *Word Embedding Revisited: A New Representation Learning and Explicit Matrix Factorization Perspective* 7.

Lu, Z., Du, P., Nie, J.-Y., 2020. VGCN-BERT: Augmenting BERT with Graph Embedding for Text Classification. arXiv:2004.05707 [cs, stat].

Mostafazadeh, N., Chambers, N., He, X., Parikh, D., Batra, D., Vanderwende, L., Kohli, P., Allen, J., 2016. *A Corpus and Evaluation Framework for Deeper Understanding of Commonsense Stories*. arXiv:1604.01696 [cs].

Mostafazadeh, N., Roth, M., Louis, A., Chambers, N., Allen, J., 2017. *LSDSem 2017 Shared Task: The Story Cloze Test*, in: Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-Level Semantics. Presented at the Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics, Association for Computational Linguistics, Valencia, Spain, pp. 46–51. <https://doi.org/10.18653/v1/W17-0906>

Reimers, N., Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv:1908.10084 [cs].

Roemmele, M., Kobayashi, S., Inoue, N., Gordon, A., 2017. *An RNN-based Binary Classifier for the Story Cloze Test*, in: Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-Level Semantics. Presented at the Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics, Association for Computational Linguistics, Valencia, Spain, pp. 74–80. <https://doi.org/10.18653/v1/W17-0911>

Schwartz, R., Sap, M., Konstas, I., Zilles, L., Choi, Y., Smith, N.A., 2017. *The Effect of Different Writing Tasks on Linguistic Style: A Case Study of the ROC Story Cloze Task*. arXiv:1702.01841 [cs].

Schwartz, R., Sap, M., Konstas, I., Zilles, L., Choi, Y., Smith, N.A., 2017. *Story Cloze Task: UW NLP System*, in: Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-Level Semantics. Presented at the Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics, Association for Computational Linguistics, Valencia, Spain, pp. 52–55. <https://doi.org/10.18653/v1/W17-0907>

Srinivasan, S., Arora, R., Riedl, M., 2018. *A Simple and Effective Approach to the Story Cloze Test*, in: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). Presented at the Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), Association for Computational Linguistics, New Orleans, Louisiana, pp. 92–96. <https://doi.org/10.18653/v1/N18-2015>

Trischler, A., Wang, T., Yuan, X., Harris, J., Sordoni, A., Bachman, P., Suleman, K., 2017. *NewsQA: A Machine Comprehension Dataset*. arXiv:1611.09830[cs].

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. *Attention is All you Need* 11

Verspoor, K., Cohen, K.B., 2013. *Natural Language Processing*, in: Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H. (Eds.), *Encyclopedia of Systems Biology*. Springer New York, New York, NY, pp. 1495–1498. https://doi.org/10.1007/978-1-4419-9863-7_158

Wang, B., Kuo, C.-C.J., 2020. SBERT-WK: A *Sentence Embedding Method by Dissecting BERT-based Word Models*. arXiv:2002.06652 [cs].

Zhou, M., Duan, N., Liu, S., Shum, H.-Y., 2020. *Progress in Neural NLP: Modeling, Learning, and Reasoning*. *Engineering* 6, 275–290. <https://doi.org/10.1016/j.eng.2019.12.014>