

REFERENCES FOR FURTHER READING

1. Dan Jurafsky and James H. Martin. Speech and Language Processing (3rd ed. draft)
2. Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing
3. Lund, K., and C. Burgess. "Hyper-space analogue to language." Behavior Research, Instruments, & Computers (1996).
4. Douglas L. T. Rohde, Laura M. Gonnerman, David C. Plaut, "An Improved Model of Semantic Similarity Based on Lexical Co-Occurrence", Communications of the ACM, 2006
5. Deerwester, Scott, et al. "Indexing by latent semantic analysis." Journal of the American Society for information science 41.6 (1990): 391-407
6. Tomás Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean, "Efficient Estimation of Word Representations in Vector Space", 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013. 1301.3781.pdf (arxiv.org)
7. Pennington, J.; Socher, R. & Manning, C. D., "GloVe: Global Vectors for Word Representation," Empirical Methods in Natural Language Processing (EMNLP), 2014, 1532-1543
8. Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). pages 2227-2237
9. Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: Neural Comput. 9.8 (Nov. 1997), pp. 1735-1780. issn: 0899-7667. doi:1162/neco.1997.9.8.1735. url: <http://dx.doi.org/10.1162/neco.1997.9.8.1735>
10. Brown, Peter F., et al. "A statistical approach to machine translation." Computational linguistics 16.2 (1990): 79-85.
11. Brown, Peter F., et al. "The mathematics of statistical machine translation: Parameter estimation." Computational linguistics 19.2 (1993): 263-311.
12. Koehn, Philipp, Franz J. Och, and Daniel Marcu. Statistical phrase-based translation. University of Southern California Marina Del Rey Information Sciences Inst, 2003.
13. Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014).
14. Bengio, Yoshua, Réjean Ducharme, and Pascal Vincent. "A neural probabilistic language model." Advances in Neural Information Processing Systems 13 (2000).
15. Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. "Effective approaches to attention-based neural machine translation." arXiv preprint arXiv:1508.04025 (2015).
16. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. 2018 Oct 11.
17. Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017)
18. Hori, Chiori, and Takaaki Hori. "End-to-end conversation modeling track in DSTC6." arXiv preprint arXiv:1706.07440 (2017).