





















- Neural Networks” , In Advances in neural information processing systems, pages 3104–3112.
- [2] Tao Ge, Furu Wei, Ming Zhou, 2018, “Fluency Boost Learning and Inference for Neural Grammatical Error Correction” , *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 1–11.
- [3] Po-Lin Chen, 2017, “Chinese Grammatical Error Diagnosis System” , Retrieved from <http://ir.lib.cyut.edu.tw:8080/handle/310901800/34228>
- [4] Taku Kudo, 2007, “CRF++ : Yet Another CRF toolkit” , <https://taku910.github.io/crfpp/>.
- [5] NLP-TEA3 CGED Shared Task, 2016, <https://www.aclweb.org/anthology/W16-4906>.
- [6] Dzmitry Bahdanau, KyungHyun Cho, Yoshua Bengio, 2014, “Neural Machine Translation by Jointly Learning to Align and Translate”, arXiv preprint arXiv:1409.0473.
- [7] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio, 2014, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling” , arXiv preprint arXiv:1412.3555v1.
- [8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13)*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Eds.), Vol. 2. Curran Associates Inc., USA, 3111-3119.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, 2018, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding ,(Submitted on 11 Oct 2018 (v1), last revised 24 May 2019 (this version, v2))
- [10] Eric Sven Ristad, Peter N. Yianilos, 1998, IEEE,” Learning String-Edit Distance”, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 20, No. 5.