Sequence with Sequence Learning with Neural Networks. Advance in Information Processing System

Khoi.Nguyen, Loc.Nguyen, Hoang.Dang, Phuong.Nguyen

0.1 Introduction

(Tsubaki, Tomii, & Sese, 2019)

Sequence-to-Sequence (seq2seq) modeling has rapidly become an important general-purpose NLP tool that has proven effective for many text-generation and sequence-labeling tasks. Seq2seq builds on deep neural language modeling and inherits its remarkable accuracy in estimating local, next-word distributions. In this work, we introduce a model and beam-search training scheme, based on the work of Daume III and Marcu (2005), that extends seq2seq to learn global sequence scores. This structured approach avoids classical biases associated with local training and unifies the training loss with the test-time usage, while preserving the proven model architecture of seq2seq and its efficient training approach. We show that our system outperforms a highly-optimized attention-based seq2seq system and other baselines on three different sequence to sequence tasks: word ordering, parsing, and machine translation.

0.2 Approach

(Sehovac & Grolinger, 2020)

0.3 Application

References

Sehovac, L., & Grolinger, K. (2020). Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention. *IEEE Access*, *8*, 36411–36426. Retrieved from `https://doi.org/10.1109/ACCESS.2020.2975738` doi: 10.1109/AC-CESS.2020.2975738

Strobelt, H., Gehrmann, S., Behrisch, M., Perer, A., Pfister, H., & Rush, A. M. (2019). Seq2seq-vis: A visual debugging tool for sequence-to-sequence models. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 353-363. doi: 10.1109/TVCG.2018.2865044

Tsubaki, M., Tomii, K., & Sese, J. (2019). Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. *Bioinform.*, *35*(2), 309–318. Retrieved from `https://doi.org/10.1093/bioinformatics/bty535` doi: 10.1093/bioinformatics/bty535