Jonathan Ho

CS 4395

Reading ACL Papers

For this assignment, we were to pick a paper from the Association of Computational Linguistics Anthology website to read and summarize. The papers that appear on the website are all peer-reviewed papers which focus on numerous different disciplines of machine learning and natural language processing. One of the papers I came across was titled “QuoteR: A Benchmark of Quote Recommendation for Writing”. As I like to quote people on occasion, the topic got me interested and I read the paper in more detail.

The authors of the paper are Fanchao Qi, Yanhui Yang, Jing Yi1, Zhili Cheng, Zhiyuan Liu, and Maosong Sun. The first four authors are affiliated with Tsinghua University in Beijing, China. The last two are associated with the Beijing National Research Center for Information Science and Technology and International Innovation Center of Tsinghua University, Shanghai, China.

The authors of this paper wanted to solve the problem about picking good quotes while writing. People commonly use quotes in their writing. However, it sometimes can be difficult to find a good quote to use that fits contextually. This is a problem that other authors have recognized as they have looked into various ways for quote recommendations.

Prior work of quote recommendation date all the way back to 2015. One aspect was integrating neural networks to be able to make a quote recommendation model. Some of these included deep learning variations such as LSTM, CNN, and GRU. Some authors have combined these variations to create a unique model, or they have used other techniques such as a learning-to-rank framework. The other authors would train and test a model for quote recommendation, but the dataset of quotes they would use is a private one not available outside of their research, which led to it being a problem for those who wanted to further research into quote recommendation.

To solve this problem of a private dataset, the authors of the paper created QuoteR (from quote **Quote R**ecommendations), which is a large, open quote recommendation dataset for English, standard Chinese, and classical Chinese. This dataset would be used to test existing quote recommendation models thoroughly and extensively. While testing these models, the authors of the paper have discovered problems with all the tested models and have come up with a novel model that would correct these problems and outperform all other models. This model would be based on BERT (Bidirectional Encoder Representations from Transformers).

To evaluate the model the authors proposed, they compared it with models serving as baselines, models from other works, and similar BERT models used for different purposes. The authors also established quantitative evaluation metrics for testing each model. After testing the models, ablation studies were carried out to make sure their training strategy for their model was effective. Human evaluators also evaluated the quotes selected to see if the quote would fit contextually in writing.

Of the six authors, only four of them had profiles for Google Scholar: Fanchao Qi, Zhili Cheng, Zhiyuan Liu, and Maosong Sun. They had 1062, 2, 30181, and 31217 citations total respectively, which would lead to Maosong Sun having the most citations for their publications.

I believe their work is significant because people quote others a lot. If a person can pick the right quote for their writing, then the quote could help their audience relate more with the author. It might even help the audience gain an understanding of what the author is trying to convey. At the very least, it would help the author to figure out a relevant quote to use instead of a bad quote or one that does not actually exist. So, while this research may not be as big as ChatGPT, this is still something that should not be taken for granted. It could even potentially be integrated into ChatGPT to help with quote recommendations.