University *of Ljubljana*
Faculty *of Computer and Information Science*

Paraphrasing sentences

Luka Boljević, Matjaž Bizjak, Mitja Vendramin

Abstract

The abstract goes here (at some point in time)

Keywords
Keyword1, Keyword2, Keyword3 ...

Advisors: Slavko Žitnik

Introduction

- The project is about paraphrasing sentences, meaning that an input sentence is converted to a different one, that still conveys the same meaning.
- The idea is to use a Slovene dataset (ccKres or ccGigafida), translate it to English, and then translate it back to Slovene to obtain paraphrases. Use multiple translators for more paraphrases?
- Manually evaluate a randomly sampled part of the created dataset (using a method or combination of methods from the 6 articles)
- Once we obtain an idea of the quality, train some generative models
- Use a test part of the dataset to evaluate the models, using automatic means (BLEU, ROUGE, BERT-iBLEU proposed by Niu et. al., etc) and the same manual evaluation method previously used
- Victory

Some useful links:

- ccKres
- ccGigafida
- Main translator - Slovene_NMT
- Auxiliary translators - One or Two (only Slovene to English though)
- Another possibility - eTranslation (requires some sort of special access)
- Translators found here

Related work

In the article Unsupervised Paraphrasing with Pretrained Language Models [1], they propose a paraphrase generation called "Dynamic Blocking". For automatic evaluation they used two different evaluations. The first one was BERT-iBLEU, which is a harmonic mean of BERTscore and one minus self-BLEU. The metric proved to be significantly better corelated to human evaluation than traditional metrics. The second type of automatic evaluation was using iBLEU, BLEU and ROUGE. Human evaluation was done by comparing this model with an already existing one, and compare paraphrases from both to see which one the evaluators liked more.

A model (DNPG) that generates paraphrases was proposed in article [2]. They tested their model on QQP (human-labeled) and WikiAnswers (automatically labeled). They used BLEU, ROUGE, iBLEU, and compared their DNPG model with 4 existing neural-based models. Human evaluation used six human assessors that each evaluate 120 groups of paraphrase candidates, given input sentences of course. Each group consists of the output paraphrases from 4 models, including the authors'. The evaluators then ranked the 4 candidates from 1 (best) to 4 (worst) based on their readability, accuracy, and surface dissimilarity to the input sentence. In Appendix B, they gave a detailed evaluation guide, and listed some criterion's for paraphrase quality.

The article A Deep Generative Framework for Paraphrase Generation [3] explains a method of paraphrase generation using a deep generative models with sequence-to-sequence models, where given an input sequence it generates a paraphrase. They use their own encoder and decoder to generate sequences which also works on paraphrases not previously learned - one sequence generates multiple phrases. The human evaluation in this article consisted of multiple human evaluators scoring some subset of results on relevance and

readability.

The article Paraphrase Generation with Deep Reinforcement Learning [4] has a different approach. They propose an approach consisting of a generator and evaluator. The generator is first trained to generate paraphrases, which are then evaluated by the evaluator. The generator is then fine-tuned by reinforcement learning where the reward is determined by the evaluator. For the evaluator learning they propose two different methods - supervised learning and inverse reinforcement learning. The human evaluation is done very similarly to article [3]. They propose two metrics - relevance and fluency.

In the article [5] they talk about retriveal-based method for paraphrase generation with Micro Edit Vectors for fine grade control over editing process. The method is composed of Retriever and Editor, the retriever selects a pair from the training corpus which is most similar to input sequence. Editor is further split to Edit provider, which computes Micro Edit Vector (MEV) and Edit performer, which rephrases the input sequence by utilizing the computed vector.The method was tested on Quora question pair dataset and the Twitter URL paraphrasing corpus. Automatic comparison was made using BLEU, ROUGE and METEOR methods, while manual testing involved six annotators.

Article [6] proposed a Syntax Guided Controlled Paraphraser (SGCP), an end-to-end framework for syntactic paraphrase generation. It's a controlled text generation that adapts the sentence using various constraints. SGCP is comprised of 3 parts. Sentence encoder, syntactic tree encoder and a syntactic-paraphrase-decoder. They used ParaNMT-small and QQP-Pos datasets. Automatic valuation was made using BLEU, METEOR, ROUGE-1, ROUGE-2, ROUGE-L, Syntactic transfer (TED-R and TED-E) and Model-based evaluation, while manual evaluation consisted of three judges,

Methods

Results

Discussion

Conclusion

Acknowledgments

References

[1] Tong Niu, Semih Yavuz, Yingbo Zhou, Nitish Shirish Keskar, Huan Wang, and Caiming Xiong. Unsupervised paraphrasing with pretrained language models, 2020.

[2] Zichao Li, Xin Jiang, Lifeng Shang, and Qun Liu. Decomposable neural paraphrase generation. *arXiv preprint arXiv:1906.09741*, 2019.

[3] Ankush Gupta, Arvind Agarwal, Prawaan Singh, and Piyush Rai. A deep generative framework for paraphrase generation, 2017.

[4] Zichao Li, Xin Jiang, Lifeng Shang, and Hang Li. Paraphrase generation with deep reinforcement learning. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3865–3878, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

[5] Amirhossein Kazemnejad, Mohammadreza Salehi, and Mahdieh Soleymani Baghshah. Paraphrase generation by learning how to edit from samples. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6010–6021, Online, July 2020. Association for Computational Linguistics.

[6] Ashutosh Kumar, Kabir Ahuja, Raghuram Vadapalli, and Partha Talukdar. Syntax-Guided Controlled Generation of Paraphrases. *Transactions of the Association for Computational Linguistics*, 8:330–345, 06 2020.