A Boring-yet-effective Approach for product reranking

Vitor Jeronymo, Guilherme Rosa, Surya Kallumadi,
Roberto Lotufo, Rodrigo Nogueira

INTRODUCTION

In this work, we detail our submission to the Amazon KDD Cup 2022 for Task 1, whose goal is to evaluate ranking methods that can be used to improve the customer experience when searching for products.

Our solution is based on the monoT5 model, that demonstrated strong effectiveness in various passage ranking tasks in different domains.

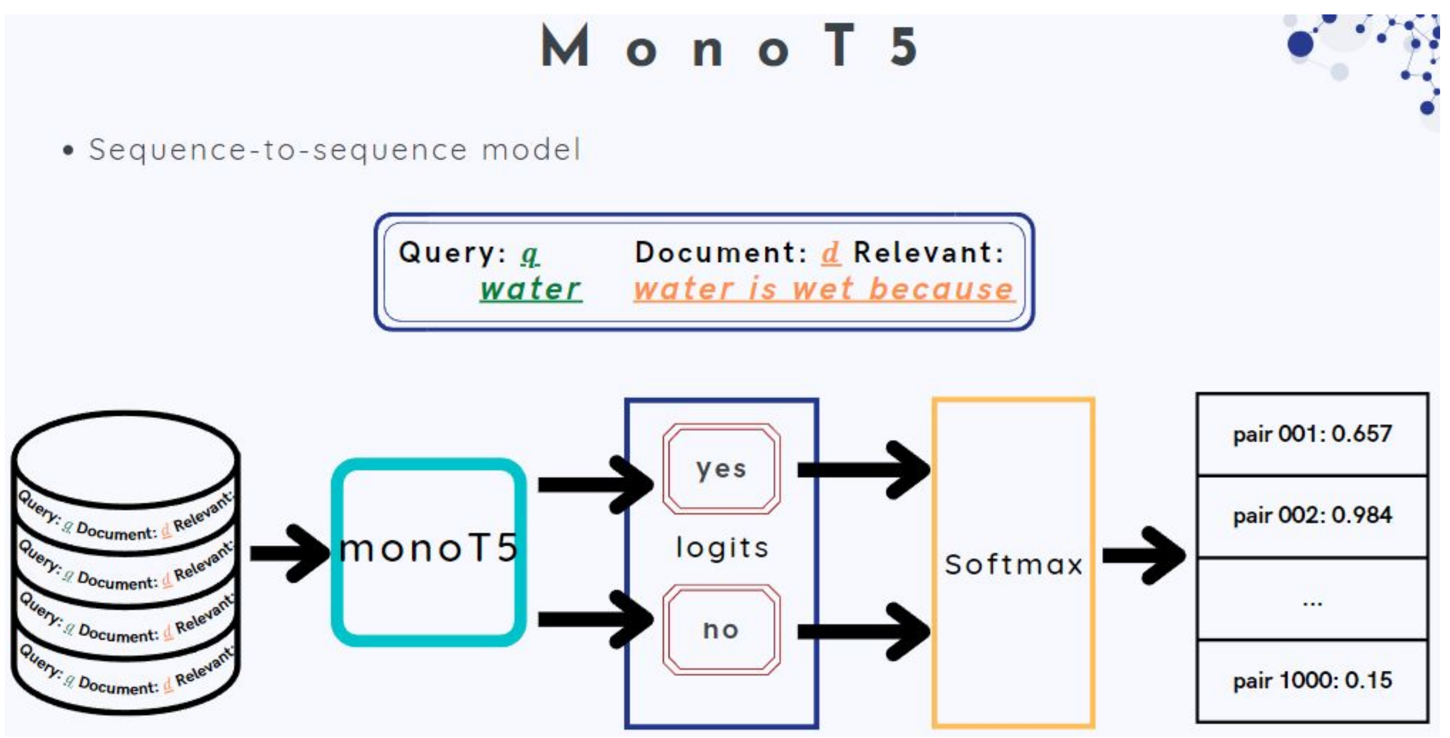

Figure 1. monoT5 reranking pipeline.

RELATED WORK

We qualify our method as “boring”, since it is well known in the recent IR literature that models with more parameters can outperform smaller ones with task-specific adaptations.

The monoT5 model is at the top of the leaderboard for the following datasets:

- TREC 2004 Robust Track
- COLIEE2021
- COLIEE2022
- TREC-COVID
- Precision Medicine 2020
- Health Misinformation 2020
- Health Misinformation 2021

METHODOLOGY

We first finetune a multilingual T5 model on the mMARCO dataset, which is the translated version of MS MARCO in 9 languages. Then, we further finetuned the model on the training data of tasks 1 and 2 of the competition.

Although the approach is “boring”, it’s not trivial because the mT5 3.7 billion parameters model training is unachievable in a regular Tesla V-100 GPU due to memory restrictions, hence we had to use the Mesh library to train it in a TPUv3, taking around 7 days for 100k steps, with batch of 128.

Products are presented to the model as the concatenation of the fields product_title, product_description, product_bullet_point, product_brand and product_color_name, joined by whitespaces. Each field is processed by the Beautiful Soup library to clean any remaining HTML tags that may appear.

As to the task labels, we considered “exact” to represent the yes token, and all other labels as no, for the mT5 classes.

Figure 2. Translated languages through Google Translate API.

RESULTS

Our model achieved an nDCG@20 of 0.9012 and 0.9007 on the public and private test sets, respectively, placing us in the ninth place on the leaderboard and only 0.0036 behind the first position.

AMAZON KDD CUP '22

TASK 1

RANK	TEAM	NDCG (PRIVATE)	NDCG (PUBLIC)
1st	www	0.9043	0.9057
2nd	qinpersevere	0.9036	0.9047
3rd	day-day-up	0.9035	0.9056
4th	GraphMIRAcles	0.9028	0.9036
5th	ZhichunRoad	0.9025	0.9035
6th	ETS-Lab	0.9014	0.9025
7th	ALONG	0.9014	0.8999
8th	ljr333	0.9008	0.9012
9th	NeuralMind	0.9007	0.9012
10th	zackchen	0.8998	0.9030

Figure 3. Final leaderboard for task 1.

CONCLUSION

Text in Calibri, 40, Bold. Text in Calibri, 40 points, Bold. Text in Calipointsbri, 40 points, Bold. Text in Calibri, 40 points, Bold. Text in Calibri, 40 points, Bold.

Acknowledgments (Calibri, 40 points, bold)

Google, for providing trial TPUsv3 for training the model

References (Calibri, 40 points, bold)

Nogueira, Rodrigo, Zhiying Jiang, and Jimmy Lin. "Document ranking with a pretrained sequence-to-sequence model." arXiv preprint arXiv:2003.06713 (2020).

Bonifacio, Luiz Henrique, et al. "mMARCO: A Multilingual Version of the MS MARCO Passage Ranking Dataset."