Digital Media Laboratory

Weekly Report September 21, 2021

Ali Jandaghi Alaee
alialaee98@gmail.com

Report No. 9
00/06/29- 00/06/23

Abstract

First of all, the Skip-Gram model for learning the embeddings of products and queries was successfully executed, and a comparison between this model and the default ranking system took place. After that, some manipulations on hyper parameters of the model were done to enhance the strength of the model.

Description

- A complete execution was done to obtain the product based version of embedding space. The details of this execution are given in the table below.

Vocabulary size	500 queries + 7267 products
Dimension of vectors	300
Number of training samples	1'662'905
Number of epochs	1

After the model learned the embeddings space of which the queries and products are related, its performance was evaluated in comparison with the default method used in the data set. The results are illustrated on the table below.

Metrics \Ranking systems	Default Model	Skip-Gram Model
Result of nDCG based on the first click on each query	27.53	18.03
First hit ratio	11.54 percent	6.53 percent

Considering the fact that the default method has provided the users with the links while the data were sampled, obviously, evaluation results of the default method are somehow biased; meaning that the performance of the Skip-Gram model is acceptable. To illustrate the bias, when a user sends a query and gets a result, the probability that the user clicks on the available results, though they are not a perfect match for their search, is relatively high.

- In addition, in three different executions, the minimum number of repetition of a training data to be preserved in the data set was changed. The goal was to get the best results in the evaluation method. Although 10, 30 and 50 were used as the minimum repetition requirement in different executions, no significant difference observed in the evaluation results.

Next Week

1. Starting the review of literature and data analyzing of the second phase of the project (Personalisation).

References