Metric learning for Finding Paraphrases.

Fedor Kitashov et al.

Abstract— **Finding paraphrases in Russian is a tricky problem due to nearly N! number of ways to construct a valid sentence out of N words. In this paper we introduce the neural pipeline to make finding paraphrases algorithms more robust. We use bi-LSTM network over Fasttext vectors to measure the similarity of two given texts. We can use the returned value to rank the texts in terms of similarity. This method outperforms tf-idf + KNN baseline in both precision and recall.**