

# RANKING WITH NEURAL NETWORK DERIVED DOCUMENT VECTORS

**Anton Serjogin**

Centre for Information and Speech Processing, LMU

`anton.serjogin@gmail.com`

15.05.2017

Motivation: come up with a ranked recommendations system that lets the user find a needed document by his input in the search query from different domains. The algorithm also filters the information by the user's metadata and clickthrough data. The Project is called IROM. The axiom is: efficient document representation. The document representation in traditional Information Retrieval is TF-IDF. The problems, however, are ignorance of word order and flawed word independence assumption. Neural Network also called Autoencoder, can predict the next word. The same approach should be done with documents, for instance, a traditional way is LSA (Latent Semantic Analysis), a big drawback of which is that the analysis requires an entire corpus, thus better done offline. The tasks are following: Train LSTM, Create API, Evaluate ranking performances, evaluate select features, search for constant shifts.