Project Proporsal

Team:
Arsenina Anna
Pinigin Arseny

Project:

implement wDTW from (https://www.aclweb.org/anthology/I17-1095.pdf) and consider different modifications of it, the algorithm can be also modified by triplet loss and ensemble of lstms (for sentences on the lower level and between them on the top) in order to use the model for semantic plagiarism detection.

We will try different approaches to test the idea of deep learning effectiveness in this task.

Dataset: https://pan.webis.de/clef14/pan14-web/plagiarism-detection.html