

On the convergence of mapping approaches for Bilingual Word Embeddings

Supervisor: Viktor Hangya, Silvia Severini

Examiner: Prof. Schuetze

Open to: Msc

Bilingual Word Embeddings (BWEs) are popular lightweight resources for cross-lingual NLP tasks. Approaches based on the iterative Procrustes refinement build good quality BWEs by iteratively improving the initial set word pairs in the training dictionary. The goal of this topic is to study the convergence of the iterative process (e.g. what word pairs are selected for training in the final iteration given different initialization) and to develop a more robust refinement process in order to improve the quality of BWEs. The project involves modifying the Python code of existing approaches.

Prerequisites: Experience in Python