

Initialization

Since, this is a non convex optimization problem, initial solution is very important in order to avoid the local optima. There are few approaches which try to use some supervision by using a seed dictionary to initialize the solution but their analysis reveals that the self-learning method gets stuck in poor local optima when the initial solution is not good enough, thus failing for smaller training dictionaries.

We use an unsupervised method to build an initial solution without the need of a seed dictionary, based on the observation that, given the similarity matrix of all words in the vocabulary, each word has a different distribution of similarity values. Two equivalent words in different languages should have a similar distribution, and we can use this fact to induce the initial set of word pairings. We combine this initialization with a more robust self-learning method, which is able to start from the weak initial solution and iteratively improve the mapping.