

Word Space Models, Local structure and Global dimensionality

Jussi Karlgren
SICS, Sweden

Abstract:

Word space models, in the sense of vector space models built on distributional data taken from texts, are used to model semantic relations between words. We argue that the high dimensionality of typical vector space models lead to unintuitive effects on modelling likeness of meaning and that the local structure of word spaces is where interesting semantic relations reside. We show that the local structure of word spaces has substantially different dimensionality and character than the global space and that this structure shows potential to be exploited for further semantic analysis using methods for local analysis of vector space structure rather than globally scoped methods typically in use today such as singular value decomposition or principal component analysis.