

Abstract

There are a lot of methods on measuring semantic similarity between sentences, monolingual alignment is one of them. Monolingual alignment is a method that align words which have similar form or similar meaning, monolingual alignment has a pipeline contains the features, in this research the monolingual alignment features are align identical word, align PPDB, align word sequences, align named entities, align content words (dependency), align content words using surrounding words (text neighbor), align stop words, align PPDB Extended, this method was developed to do an alignment on a pair of sentences that has to be aligned but it can be used to measure semantic similarity between sentences.

The result of this method can give 91,1% of F1 for alignment dataset it's less 0,6% compared to Sultan et al. [1] which get the best result of alignment on MSR data and 64% is an average of correlation on measuring semantic similarity on dataset STS, this system got 8th rank on STS 2013, 13th rank on STS 2013, and 56th on STS 2015.

Keywords: Monolingual alignment, STS, Semantic, PPDB.