

## Abstrak

Untuk mengukur kesamaan semantik antarkalimat ada banyak metode yang dapat digunakan, salah satunya adalah *monolingual alignment*. *Monolingual alignment* adalah metode yang melakukan penyejajaran kata-kata yang memiliki kesamaan secara huruf atau kesamaan secara makna di antara dua kalimat. Dalam *monolingual alignment* ada yang disebut *alignment pipeline* yang berisi penggunaan urutan fitur-fitur, fitur yang digunakan dalam penelitian ini yaitu *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*. Metode ini dikembangkan untuk melakukan *alignment* terhadap data pasangan kalimat *alignment* tetapi dapat digunakan untuk mengukur kesamaan semantik antarkalimat. Penggunaan metode *monolingual alignment* dengan fitur *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* dapat menghasilkan nilai F1 91,1% untuk data MSR (*alignment*) berbeda 0,6% dari aligner Sultan et al. [1] yang saat ini memiliki nilai tertinggi dan 64% nilai rata-rata korelasi untuk dataset STS. Sistem ini menempati peringkat 8 untuk STS 2013, 13 untuk STS 2014, dan 56 untuk STS 2015.

**Kata kunci:** *monolingual alignment*, STS, semantik, MSR, PPDB.