

Abstrak

Bahasa memiliki beragam bentuk. Hal tersebut akan mempengaruhi bentuk kalimat yang dibuat. Dengan keberagaman tersebut terdapat kemungkinan beberapa kalimat mempunyai kata-kata yang sebenarnya memiliki makna yang sama. Oleh karena itu penelitian ini merancang interpretable semantic textual similiarity untuk mendapatkan skor dan label kesamaan. Untuk mendapatkan skor tersebut, digunakan IXA Pipeline untuk mendapatkan data chunking otomatis dari teks yang kemudian diukur juga skor akurasi dengan evaluation metrics. Dari penelitian yang telah dilakukan, didapatkan F1 Measure dengan 0.2638 untuk F1 Alignment, 0.1703 untuk F1 Type, 0.2418 untuk F1 Score, 0.1703 untuk F1 Type+Score

Kata kunci : Interpretable Textual Similarity, IXA Pipeline, evaluation metrics, chunk