

INTERPRETABLE STS BERBAHASA INGGRIS DENGAN METODE CHUNKING OTOMATIS, CHUNK ALIGNMENT DAN SEMANTIC RELATION PREDICTION

Nasib Setia¹, Moch. Arif Bijaksana²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹setianasib@students.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id,

Abstrak

Penelitian dibidang Linguistik merupakan salah satu penelitian untuk mengukur tingkat kesamaan makna antara dua buah teks dengan memberikan label keterhubungan dan skor kesamaan antar teks berdasarkan proses alignment. Salah satunya dengan Interpretable STS (Semantic Textual Simillarity) untuk mengukur tingkat kesamaan dua buah teks. Tujuan dari tugas akhir ini adalah untuk membuat program menghitung tingkat kesamaan makna pasangan kata dalam bahasa Inggris dengan cara pemberian label keterhubungan kemudian skor kesamaan antar teks berdasarkan proses alignment dengan metode chunk otomatis, chunk alignment. Tahapan yang dilakukan dalam melakukan penelitian ini adalah mengumpulkan data set pasangan kata dalam bahasa Inggris, kemudian dengan proses chunking otomatis kata akan di potong-potong dan kemudian masuk dalam proses alignment. Selanjutnya akan diberikan label keterhubungan antar potongan kata, dan kemudian tahapan terakhir dengan memberikan skor kesamaan antara potongan kata. Dengan penelitian ini diharapkan dapat mengetahui mengenai tingkat kesamaan makna pasangan potongan kata bahasa Inggris dilihat dari nilai kesamaan makna pasangan potongan kata berbahasa Inggris tersebut melalui nilai kesamaan dan label keterhubungan. Kata Kunci : Interpretable STS, Semantic Textual Simillarity, Chunking Otomatis, Chunk Alignment, Berbahasa Inggris.

Abstract

Research in the field of Linguistics is one of the research to measure the level of similarity between two pieces of text with the labeling connectedness and similarity scores between texts based on alignment process. One of them with Interpretable STS (Semantic Textual Simillarity) to measure the level of similarity of two pieces of text. The purpose of this final project is to make the program to calculate the level of equality of the word pair in English by way of labeling connectedness then score of similarity between text based on alignment process with automatic chunk method, chunk alignment. Stages performed in conducting this research is to collect data set of word pairs in English, then by chunking process word will automatically cut into pieces and then entered in the process of alignment. Next will be labeled connectivity between pieces of the word, and then the last stage by giving a score of similarities between the pieces of the word. With this research is expected to know about the level of similarity of the meaning of the English word piece pair seen from the value of similarity of meaning of the English word piece pair through the value of similarity and label connectedness. Keywords: Interpretable STS, Semantic Textual Simillarity, Automatic Chunking, Chunk Alignment, Speaking English
