Abstract

Research in the field of Linguistics has an appeal that is large enough at the moment, one of which is the research interpretable STS (Semantic Textual Simillarity) which is a study to measure the level of similarity of meaning between the two create a text by labeling of connectedness and similarity score between the text based alignment process. The purpose of this final project is a program to calculate the degree of similarity of meaning between a couple verses of the Koran in English translation Saheeh international version with Chunking process is automated. To conduct this research, created the system in the form of unsupervised system. An outline of the steps being taken in this study is to collect data translation couple verses of the Koran, then the data through a process Chunking system automatically and entered in the alignment process. Furthermore, given the label connectivity between pieces of verses of the Koran, and the last stage of the calculation score similarities between pieces of verses of the Koran. Their research is expected to determine optimally on the similarity of meaning translations of the Koran through a couple pieces of similarity values and connectedness owned label. This study has an accuracy value for chunking system by 88% and the value of F_1 Ali (F_1 for alignment) which is the highest at 92.9% .

Keywords: Semantic Textual Similarity, Interpretable STS, unsupervised system, Chunking, alignment