

TELKOM UNIVERISTY

Abstract

School of Computing

The Graduate School

**Answer Selection Using Word Alignment Based On Part Of Speech Tagging In
Community Question Answering**

by Ade Sutedi

Community Question-Answering (CQA) is one of online forums where user allowed to ask question and the other users can reply or answer the related question or problems. Due to CQA has no restrictions in conveying questions or answers, there are comments that are not appropriate with the problems. To solve these problems, combining lexical and semantic features has been developed in the previous research. But, the approach more adequate for similarity task rather than question answering. According to this problem, there are several problems that can be enhanced. First, vector representation counts exactly matched words, it does not effective to cover other words that have relatedness between two pairing words. Second, noun overlap for similarity measure in pairing words can't define that the two words are similar. So, it must be defined that the pairing POS tag is the same meaning or relatedness. In this study, unsupervised lexical and semantic similarity method employed with a different approach from the previous method in verbatim and contextual similarities. The data was taken from SemEval 2017 competition which focuses on Question-Answer Similarity task. The experiment result for precision (Mean Average Precision) score shows the significant improvement from 0.6742 to 0.6845, 1.03 % higher than previous research in CQA. This improvement comes from lexical similarity, which is not just from noun pattern but also taken from verb pattern. Furthermore, semantic similarity has an important role in determining which words that have the same pattern and meaning to define relevancy between them.

Keywords: Community Question-Answering, Unsupervised, Lexical and Semantic Similarity, Relevancy.