

ABSTRACT

Essay is a response from a student against a complex learning process where many researcher are still using the answer essays to test student comprehension level. Computer that become important component in the learning process easier in doing the assessment in multiple choice exam and a short field accurately compared to the essay exam because the answer that there must be the same, either by choice or by his words. Whereas type multiple-choice exam that is lacking in terms of the learner at the level of understanding about the exam.

This final assignment implements Probabilistic Latent Semantic Analysis (PLSA) that can identify the important sentences are semantically useful for generating output a ekstractive assessment. On the application of this system, the previous document should be in preprocessing, then extracted into terms and proceed with stopword removal and stemming process. In providing of output scoring, the documents that has been modeled using PLSA with latent variable parameter then matched using cosine similarity and euclidean distance technique.

Once the testing automatic essay grading system is done using 2 test data, it can be concluded that the smaller scale value that assigned by lecturer make the average accuracy rate higher, with an average accuracy 53 – 77% for data 1 and 82 – 91% for data 2 in which data 1 has a value scale manually by lecturer 0 – 19 and data 2 has been scale manually 0 – 10.

Keywords: *PLSA, preprocessing, automated essay grading, stopword removal, stemming, latent variable, cosine similarity, euclidean distance.*