

Sentiment Analysis Against Student Feedback About the Language Center Service (Case Study : Telkom University)

Adi Satria Pangestu¹, Yuliant Sibaroni², Mohamad Syahrul Mubarak³

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

¹adisatriaa@students.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id,

³msyahrulmubarak@telkomuniversity.ac.id

Abstract

Language center services are language services that serve the needs of the students involved in linguistics, including ITP TOEFL, ECCT, EPrT, and English language courses. Students really need the value of the English language test to meet the requirements at the level of study. Quality of service at the service center language needs to be considered in order to provide the best for its users. To improve the level of service, it is necessary to evaluate the service by knowing the comments and responses given, then do a sentiment analysis of the response data. Sentiment analysis includes a branch of text mining research that is useful for classifying text documents in the form of sentiment based opinions. This study aims to calculate the magnitude of positive and negative sentiments towards language center services. In addition to looking at the sentiments, this study also aims to see the results of evaluation using stemming and without stemming at the preprocessing stage. This study uses K-Nearest Neighbor (KNN) and Multinomial Naïve Bayes (MNB). The feature extraction used for both classifications is Term Frequency-Inverse Document Frequency (TF-IDF). The use of data in this study is 5000 responses and uses cross validation evaluation. The pre-processing stages carried out in this study consisted of tokenizing, cleansing, case folding, stopword removal, and stemming. System accuracy results obtained in the KNN method with $k = 3$ produce 63.1% accuracy value using stemming, while without stemming produces 64.84%. In MNB classification produces accuracy value of 84.63% using stemming, while without stemming is 85.62%. So that the classification of MNB is better than KNN, and without using stemming produces better accuracy than using it. Then the sentiments taken from the best classification namely naïve bayes multinomial without stemming gave the result that sentiment on central service data was more dominant to get positive sentiment of 84%, while negative sentiment was only 16%.

Keywords: sentiment analysis, k-nearest neighbor, naïve bayes, Information System Services.