ABSTRACT

Technology is currently developing rapidly in the digital direction. advances in technology, computers and telecommunications have supported the development of internet technology. Bukalapak ranks third in the top 10 Indonesian E-commerce issued by Iprice Insight. With this rating, Bukalapak can improve the quality and quantity of its services by knowing the results of customer reviews so that later the services provided to customers will be better and can improve ratings. There are too many reviews from Bukalapak customers, so it is difficult and takes a long time to classify and analyze customer reviews if it is done manually. Therefore, we need a method that can classify customer reviews. The method used to classify reviews is the Support Vector Machine. The reviews will be classified into two types, namely positive reviews and negative reviews. The stages for classifying in this study are data preprocessing, feature extraction into TF-IDF, SVM analysis, and evaluation. There are 3 scenarios used in this study, namely comparison testing 60:40, 70:30, and 80:20 training. The results of the classification with SVM and linear kernel functions on the training data show that the three ratios have accuracy above 80%, namely 89.18% for the ratio 60:40, 89.14% for the ratio 70:30, 89.06% for the ratio 80:20. Of the three ratios, the best model formed by SVM is a ratio of 60:40. The evaluation of the best model from SVM obtained an accuracy of 85%, Recall of 79%, Precision of 89%, and F1-Score of 84%. The results of the K-Fold Cross Validation with 10 Folds show results that are not much different from the evaluation, which is an average of 84%. The results of the positive category classification can be used as a reference to maintain service quality and the negative classification results can be used as evaluation material in improving services at Bukalapak.

***Keywords: Customer Reviews, Classification, SVM, Linear kerne*l**