ABSTRACT

Review data on the PLN Mobile application can be used as a reference to evaluate the performance of PT PLN (Persero). However, the large amount of review data on the PLN Mobile application on the Play Store platform makes manual data processing inaccurate and takes a relatively long time. With this problem, a tool is needed to make it easier for PT PLN (Persero) to process large amounts of review data effectively and in a shorter time. This final project aims to make it easier for companies to process big data on PLN Mobile application reviews on the Play Store, gain business insights for PT PLN (Persero), and get the best algorithm to perform sentiment analysis on PLN Mobile application reviews. The Machine Learning methods used to analyze Big Data are the *Transformers* Indo-BERT algorithm, *Naïve Bayes*, LSTM (Long Short-Term Memory), and *Logistic Regression*. To get the best algorithm for this sentiment analysis, the results are compared with the Accuracy values of the three algorithms. This study also develops a *topic modeling* system to find out and classify topics in a sentence, so that it can make it easier for companies to understand user opinions. This study uses the LDA (Latent Dirichlet Allocation) model to predict topics in the PLN Mobile application review.

In the *Transformers* Indo-BERT, *Naïve Bayes*, LSTM (Long Short-Term Memory), and *Logistic Regression* algorithms, the highest Accuracy value was obtained, namely 96% in the *Transformers* Indo-BERT algorithm with a Precision value of 87%, Recall 85%, F1-Score 86%, so that the *Transformers* Indo-BERT algorithm is the recommended algorithm to PT PLN (Persero) to conduct sentiment analysis on PLN Mobile application reviews. Then in *topic modeling*, it successfully predicted 4 topics in each review, namely the topics "TRANSACTION", "APPLICATION", "SERVICE", and "COMPLAINT". From the four topics, there are keywords from each topic, sentiment labels, and review dates, so that they can be processed again into a comparison of monthly trend sentiments that can be used as business insights.

Keywords: *Big Data, LSTM, LDA, Logistic Regression, Naïve Bayes, Sentiment analyst, Transformers, Topic modeling.*