

## **ABSTRACT**

Rexzy Tarnando, 92217024

“COMPARISON OF PERFORMANCE LEVELS OF NAÏVE BAYES, K-NEAREST NEIGHBOR AND LOGISTIC REGRESSION ALGORITHMS IN TEXT CLASSIFICATION OF PEOPLE SENTIMENTS ON TWITTER (CASE: TRANSJAKARTA)”.

Nowadays social media is one to express things that are thought and felt by the community. One of the things that's much talked about is responses from the consumer of products or services. This is very useful for companies to find out the level of satisfaction of their products or services. Twitter is one of the most widely used social media by users. With this fact, it's really interesting for companies to use the data on Twitter for the company's progress generally in customer relations. In this study an analysis of public sentiments towards the use of Transjakarta. This study divides community sentiments into three classes, positive, neutral and negative. For data taken from Twitter with the results of research from June to July 2019 by dividing the data into training data and testing data. The amount of training data is 144 tweets and testing data are 36 tweets. Then for the text classification uses 3 algorithms, namely naïve bayes, k-nearest neighbor and logistic regression. Then after the results are obtained, next is to compare the performance levels of three methods by finding the highest f1 score value using micro average formula. Micro average is chosen because it's the best for calculating imbalanced datasets. The results show the naïve bayes method has the best f1 score with 0.861 value. For the next largest f1 score value is the logistic regression method with an f1 score value of 0.833, and the last is the k-nearest neighbor method with an f1 score value of 0.806.

Keyword: Naïve Bayes, K-Nearest Neighbor, Logistic Regression, F1 Score, Sentiment Analysis, Transjakarta  
(xii + 68 + Appendix)

Bibliography (1994 - 2019)