ABSTRACT

Twitter is a social media platform that is a place for many people to be able to upload various things, including uploads that contain a security threat to a system. Of course this is a dangerous thing if someone uploads a system security vulnerability. System threats that are published can be misused by others to the detriment the system owner.

To anticipate this, a system was created to detect tweets that contain threats and system vulnerabilities on Twitter. This system applies a text processing algorithm that uses the Naïve Bayes method and TF-IDF (Term Frequency – Inverse Document Frequency). This method was chosen because it is considered as one of the most effective and has good accuracy.

In this final project, the final result obtained is that the system can distinguish tweets that contain threat or vulnerability, and those that do not. With the ratio of the distribution of the dataset into training data and testing data is 70%:30% and 80%:20%, both obtained an accuracy value of 88%, a precision value of 88%, a recall of 88%, and an F1 score of 88%.

Keywords: *text mining, Naïve Bayes, TF-IDF, threats, vulnerabilities, text classification.*