

ABSTRACT

In Indonesia, Twitter is the preferred social media platform used by some to find and share information. Harassment of people or groups based on race, religion, ethnicity, nationality, sex, ability, sexual orientation, or gender identity is known as hate speech. Hate speech is often emotional and spreads through social media, especially Twitter. Hate speech is modeled in this study using Naive Bayesian models, namely Multinomial, Bernoulli and Gaussian Naive Bayes Models. The purpose of this research is to obtain a method with the highest accuracy value in the analysis of hate speech. The method used in this research is to process all data from Twitter social media and then categorize it based on the categorization of HS and non-HS sentiment categories using Multinomial Naive Bayes, Gaussian Naive Bayes, and Bernoulli Naive Bayes models. The Multinomial Naive Bayes model has the highest accuracy rate, which is 82.13%, based on the results of testing the data.

Keywords: Naïve Bayes, Multinomial, Gaussian, Bernoulli, Hate Speech