UNIVERSITAS
GADJAH MADA
Hate Speech Detection in Indonesian Using Complement Naive Bayes
ADIVA VIDYANI ERIZA, Nur Rokhman, S.Si., M.Kom, Dr.
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Hate Speech Detection in Indonesian Using Complement Naïve Bayes

Adiva Vidyani Eriza 17/408353/PA/17660

Hate speech has become a widespread problem as the use of social media continues to increase in Indonesia today. To quickly tackle this problem, the automation of detecting the hate speech is needed. The problem in hate speech detection is that there is way less data of hate speech compared to non-hate-speech, hence higher chance of creating an imbalanced dataset that can cause misclassification. Complement Naive Bayes classifier is often used to solve these cases of imbalanced dataset.

Complement Naïve Bayes was invented with the purpose of solving the problem when using the Multinomial Naïve Bayes, in which it does not perform well in the case of imbalanced dataset. Therefore in this research, the performance of Complement Naïve Bayes classifier is compared to the performance of Multinomial Naïve Bayes classifier in classifying imbalanced hate speech data, and the models are validated using the K-Fold Cross Validation.

In this experiment, despite the 1% lower accuracy score compared to the Multinomial Naïve Bayes classifier, the Complement Naïve Bayes classifier scored the highest of 89% in average recall score, that is 25% higher compared to Multinomial Naïve Bayes classifier. This indicates that Complement Naïve Bayes classifier can classify hate speech better, and has a lower chance of letting the actual hate speech go undetected.

Keywords: Complement Naïve Bayes, Hate Speech Detection, Natural Language Processing, Classification