

## ABSTRAK

### PENGARUH UNIGRAM DAN BIGRAM DALAM IMPLEMENTASI MULTINOMIAL NAÏVE BAYES UNTUK ANALISIS SENTIMEN ULASAN APLIKASI GOJEK DI GOOGLE PLAY STORE

Adrian Dwinanda Adyatma (adrianadyatma21@gmail.com)

Pada sistem klasifikasi sentimen yang menggunakan *Naïve Bayes Classifier*, metode ekstraksi fitur yang banyak digunakan adalah TF-IDF dengan n-gram 1 dan 2. Pada penelitian ini, diteliti pengaruh dan peningkatan performansi antara sistem klasifikasi yang menggunakan unigram atau bigram saja, dengan sistem yang menggunakan gabungan keduanya. Dengan 1000 data ulasan *rating* 1 (negatif) dan 5 (positif) pengguna *Gojek* dari *Google Play Store*, dan validasi performa dengan *K-Fold* di angka  $K=10$ , didapatkan sistem yang menggunakan ekstraksi fitur TF-IDF gabungan unigram dan bigram menghasilkan performansi paling baik dari ketiga sistem dengan akurasi 0.84, namun peningkatan tersebut cukup kecil dibandingkan dengan sistem yang menggunakan unigram saja di angka 0.83.

**Kata Kunci:** *Naïve Bayes, N-Gram, Analisis Sentimen, Google Play Store, TF-IDF*

## **ABSTRACT**

### ***THE EFFECT OF UNIGRAM AND BIGRAM ON THE APPLICATION OF MULTINOMIAL NAÏVE BAYES FOR SENTIMENT ANALYSIS OF GOJEK APPLICATION REVIEWS ON GOOGLE PLAY STORE***

*Adrian Dwinanda Adyatma (adrianadyatma21@gmail.com)*

*In sentiment classification systems that use Naïve Bayes Classifier, a commonly used feature extraction method is TF-IDF with n-grams of 1 and 2. In this research, the impact and performance improvement between classification systems using unigrams or bigrams alone and those using a combination of both are studied. Using 1000 data of reviews with ratings 1 (negative) and 5 (positive) from Gojek users on the Google Play Store, and performing performance validation with K-Fold at K=10, the system that uses the combined TF-IDF feature extraction of unigrams and bigrams achieves the best performance among the three systems with an accuracy of 0.84. However, the improvement is relatively small compared to the system that uses unigrams alone at 0.83.*

**Keywords:** *Naïve Bayes, N-Gram, Sentiment Analysis, Google Play Store, TF-IDF*