

PENERAPAN N-GRAM TERHADAP ANALISIS SENTIMEN PADA FITUR BELANJA TIKTOKSHOP MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

Riska Dwi Ayu Lestari¹, Bagus Setya Rintyarna², Moh. Dasuki³

Program Studi Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

e-mail: riskadwiayu10@gmail.com¹

ABSTRAK

Di era modern ini, telah banyak kita temui perusahaan komersial yang memasarkan produknya melalui *platform* sosial media. Bahkan baru ini aplikasi sosial media penyedia video berdurasi pendek atau dikenal dengan TikTok telah menyediakan fitur *e-commerce*. Tetapi tidak sedikit tanggapan dari masyarakat mengenai kehadiran fitur belanja TikTokShop ini yang diungkapkan melalui media sosial salah satunya twitter. Sehingga penelitian ini menggunakan data twitter dengan pencarian *keyword* “TikTokShop” yang akan diimplementasikan ke arah sentimen analisis menggunakan algoritma *K-Nearest Neighbor* dengan metode *cosine similarity* dalam menghitung jarak tetangga terdekat. Skenario uji yang digunakan menggunakan *10-fold cross validation*. Penelitian ini juga menggunakan fitur seleksi N-Gram yaitu *unigram*, *bigram*, dan *trigram*, dan untuk menangani data yang tidak seimbang menggunakan teknik *undersampling*. Hasil pemodelan didapatkan bahwa penerapan N-Gram terbaik yaitu model seleksi fitur *unigram*, dengan hasil akurasi nilai tetangga optimal terjadi pada nilai tetangga $k=5$ hal ini berdasarkan nilai rata-rata akurasi sebesar 89,92%, presisi sebesar 90,54% dan *recall* sebesar 87,37%. Berdasarkan hasil pengujian data *undersampling* lebih rendah daripada data sebelum *undersampling*.

Kata Kunci: Analisis Sentimen, TikTokShop, K-NN, N-Gram, *Undersampling*

**APPLICATION OF N-GRAM ON K-NEAREST NEIGHBOR
ALGORITHM TO SENTIMENT ANALYSIS OF TIKTOKSHOP
SHOPPING FEATURES**

Riska Dwi Ayu Lestari¹, Bagus Setya Rintyarna², Moh. Dasuki³
Informatics Engineering Study Program, Faculty of Engineering
University of Muhammadiyah Jember
e-mail: riskadwiayu10@gmail.com¹

ABSTRACT

In this modern era, we have encountered many commercial companies that market their products through social media platforms. Even recently, the social media application that provides short-form videos, known as TikTok, has provided e-commerce features. But not a few responses from the public regarding the presence of the TikTokShop shopping feature were expressed through social media, one of which is Twitter. So this study uses Twitter data by searching for the keyword "TikTokShop" which will be implemented towards sentiment analysis using the K-Nearest Neighbor algorithm with cosine similarity nearest neighbor distance method. The test scenario used is 10-fold cross validation. This study also uses the N-Gram selection features, namely unigrams, bigrams, and trigrams, and to handle unbalanced data using undersampling techniques. The modeling results show that the best application of N-Gram is the unigram feature selection model, with the result that the accuracy of the optimal neighbor value occurs at the neighbor value $k = 5$, this is based on an average value of accuracy of 89.92%, precision of 90.54% and recall of 87.37%. Based on the test results, the undersampling data is lower than the data before undersampling.

Keywords: *Sentiment Analysis, TikTokShop, K-NN, N-Gram, Undersampling*