

INTISARI

PERBANDINGAN *SENTIMENT ANALYSIS* PADA AKUN *INSTAGRAM* MENGGUNAKAN PEMBOBOTAN *TERM FREQUENCY* *INVERSE DOCUMENT FREQUENC (TF-IDF)* DAN *NAÏVE BAYES* *CLASSIFIER*

Oleh

Ghina Indah Fithriyah
12/330473/PA/14385

Dalam era ini internet menyebabkan peningkatan jumlah pengguna jejaring sosial. *Instagram* salah satu *platform* yang populer dalam masa kini. Dalam perkembangannya, *Instagram* kini memiliki banyak komunitas, diantaranya komunitas *beauty*, *travel*, *automotive*, *fashion* hingga *food*. Komunitas yang paling pesat perkembangannya di Indonesia adalah komunitas pencinta makanan (*food*). Mereka berkumpul dan menghadirkan para *influencer* yang sering kita sebut sebagai *instafood*, dan mereka akan membagikan informasi mengenai menu rekomendasi dari *food stand* sampai restoran sehingga dapat mempengaruhi bisnis makanan itu sendiri.

Pada penelitian ini dilakukan proses *sentiment analysis* menggunakan *Multinomial Naïve Bayes Classifier* dan *Naïve Bayes Classifier* pada data komentar dari dua akun *instafood* yaitu *kulineryogya* dan *jogjaculinary*. Data ini menggunakan sampel komentar *post* dari bulan Januari 2019 hingga Maret 2019 dan diambil sebanyak 5807 komentar. Setelah dilakukan *text processing* didapatkan jumlah *vector* dari fitur kedua akun tersebut sebanyak 3097 kata.

Pengujian dalam dua akun tersebut menggunakan *Multinomial Naïve Bayes* dan *Naïve Bayes Classifier* sebagai *classifier* dan metode validasi *Cross Validation*. Dari hasil pengujian ini didapatkan akurasi sebesar 0,9006 pada akun *jogjaculinary* dan 0,9424 pada akun *kulineryogya*. Hasil klasifikasi dengan menggunakan *Naïve Bayes Classifier* lebih tinggi dibandingkan dengan menggunakan *Multinomial Naïve Bayes*. Pada *Multinomial Naïve Bayes* bahwa nilai sentimen positif untuk akun *jogjaculinary* sebesar 97,23%, lebih tinggi jika dibandingkan dengan akun *kulineryogya* yang hanya sebesar 91,14%. Pada *Naïve Bayes Classifier* bahwa nilai sentimen positif untuk akun *kulineryogya* sebesar 99,08%, lebih tinggi jika dibandingkan dengan akun *jogjaculinary* yang hanya sebesar 97,71%. Perbedaan ini tidak signifikan karena kedua akun tersebut mengekspose niche yang sama di kota yang sama.

Kata Kunci : *sentiment analysis, naïve bayes, multinomial naïve bayes, tf-idf*

ABSTRACT

COMPARISON OF SENTIMENT ANALYSIS ON INSTAGRAM ACCOUNTS USING TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) WEIGHTING AND NAÏVE BAYES CLASSIFIER (NBC)

By

Ghina Indah Fithriyah
12/330473/PA/14385

Era of the Internet caused an increase of social media users. Instagram is one of the most popular platforms of social media. Nowadays, Instagram now has many community, including beauty, travel, automotive, fashion, and food. Food community is the fastest developing community in Indonesia. They are gathering and inviting the influencers that we often refer as an instafood. They will share about recommended food from foods stand to restaurant that improve the food business itself.

This research studying about sentiment analysis process was carried out using the Multinomial Naïve Bayes classifier dan Naïve Bayes Classifier on comments data from two instafood account named kulineryogya and jogjaculinary. The data uses the comments from the post on January 2019 to March 2019 as sample and is taken as many as 5807 comments. The result of calculating vector features from the two accounts were 3097 words after going through the text processing.

Multinomial Naïve Bayes and Naïve Bayes Classifier as classifier and Cross validation method was used to tested in these two accounts uses. From the results of this test we kows that accuracy 0.9006 on the jogjaculinary account and 0.9424 on the kulineryogya account. The classification results using Naïve Bayes Classifier was higher than the Multinomial Naïve Bayes. In Multinomial Naïve Bayes, positive sentiment value for the jogjaculinary account was 91.23%, it was higher than the kulineryogya account which was only 91.14%. In Naïve Bayes Classifier, positive sentiment value for the kulineryogya account was 99.08%, it was higher than the jogjaculinary account which was only 97.71%. The difference is not significant because both accounts are on the same niche in the same city.

Keyword : *sentiment analysis, naïve bayes, multinomial naïve bayes, tf-idf, naïve bayes classifier*