

ABSTRAK

Sepak bola merupakan olahraga paling populer di Indonesia, dengan Tim Nasional (Timnas) sebagai simbol kebanggaan nasional. Salah satu isu yang memicu perdebatan publik adalah kebijakan naturalisasi pemain asing untuk memperkuat Timnas. Opini publik terkait isu ini banyak ditemukan di media sosial, terutama Instagram, yang memiliki lebih dari 116 juta pengguna aktif di Indonesia. Untuk mengetahui sentimen ekspresi pendapat masyarakat tentang berbagai topik, termasuk isu-isu terkait sepak bola dan Timnas digunakan analisis sentimen. Penelitian ini bertujuan untuk menganalisis sentimen dari 4.258 komentar Instagram terhadap isu naturalisasi dengan membandingkan performa tiga algoritma *Naïve Bayes*: *Multinomial*, *Bernoulli*, dan *Gaussian*. Proses analisis meliputi *preprocessing*, penerjemahan data menggunakan *Google Translate API*, pelabelan dengan *VADER*, pembobotan *TF-IDF*, serta penyeimbangan data menggunakan *SMOTE*. Evaluasi dilakukan menggunakan *confusion matrix* berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil menunjukkan bahwa *Multinomial Naïve Bayes* memberikan performa terbaik dengan akurasi 74,65%, presisi 75,96%, *recall* 74,65%, dan *F1-score* 74,32% pada skenario *K-Fold* 10 dan *alpha* 0,01. Algoritma ini mengungguli *Bernoulli Naïve Bayes* dan *Gaussian Naïve Bayes*. Namun, *Gaussian Naïve Bayes* memiliki waktu komputasi tercepat (0,02 detik), disusul oleh *Multinomial* (0,03 detik) dan *Bernoulli* (0,94 detik).

Kata kunci: Analisis sentimen, Instagram, *Multinomial Naïve Bayes*, *Bernoulli Naïve Bayes*, *Gaussian Naïve Bayes*.

ABSTRACT

Football is the most popular sport in Indonesia, with the National Team (Timnas) serving as a symbol of national pride. One of the most debated issues is the naturalization of foreign players to strengthen the team. Public opinions on this topic are widely expressed on social media, especially Instagram, which has over 116 million active users in Indonesia. This study aims to analyze sentiment from 4,258 Instagram comments related to the naturalization issue by comparing the performance of three *Naïve Bayes* algorithms: *Multinomial*, *Bernoulli*, and *Gaussian*. The analysis process includes *preprocessing*, translation using the *Google Translate API*, sentiment labeling with *VADER*, *TF-IDF* vectorization, and data balancing using *SMOTE*. The classification performance is evaluated using a *confusion matrix* with metrics such as *accuracy*, *precision*, *recall*, and *F1-score*. The results show that *Multinomial Naïve Bayes* achieves the best performance with 74.65% *accuracy*, 75.96% *precision*, 74.65% *recall*, and 74.32% *F1-score* under 10-fold *cross-validation* and *alpha* 0.01. This algorithm outperforms *Bernoulli* and *Gaussian Naïve Bayes*. However, *Gaussian Naïve Bayes* has the fastest computation time (0.02 seconds), followed by *Multinomial* (0.03 seconds) and *Bernoulli* (0.94 seconds).

Keyword: Analisis sentimen, Instagram, *Multinomial Naïve Bayes*, *Bernoulli Naïve Bayes*, *Gaussian Naïve Bayes*