

## ABSTRAK

Utama, Singgih. 2026. Analisis Sentimen pada *Twitter* Pasca Dibataalkannya Pagelaran Piala Dunia U-20 yang Diselenggarakan di Indonesia Menggunakan Metode *Multinomial Naïve Bayes*. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

**Pembimbing:** Moh. Dasuki, S.Kom., M.Kom.; Habibatul Azizah Al-Faruq, M.Pd.

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap pembatalan penyelenggaraan Piala Dunia U-20 di Indonesia menggunakan metode *Multinomial Naïve Bayes*. Sebanyak 921 data komentar *Twitter* dikumpulkan melalui proses *crawling* dan dikelompokkan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Dataset awal menunjukkan ketidakseimbangan kelas yang signifikan, yaitu 859 data positif, 55 data negatif, dan 7 data netral. Oleh karena itu, diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE) untuk menyeimbangkan jumlah data menjadi 429 data pada setiap kelas. Selanjutnya, data teks diproses melalui tahapan *preprocessing* yang meliputi *cleansing*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, kemudian dilakukan pembobotan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Hasil pengujian menggunakan *K-Fold Cross Validation* menunjukkan bahwa penerapan SMOTE mampu meningkatkan performa model secara signifikan. Pada pengujian data seimbang, performa terbaik diperoleh pada *K-Fold* 9 dengan nilai akurasi sebesar 97,20%, presisi 97,31%, dan *recall* 97,22%. Berdasarkan hasil tersebut, metode *Multinomial Naïve Bayes* yang dikombinasikan dengan pembobotan TF-IDF dan SMOTE terbukti efektif dalam melakukan klasifikasi sentimen pada data yang tidak seimbang.

**Kata Kunci:** Analisis Sentimen, *Multinomial Naïve Bayes*, SMOTE, TF-IDF, *Twitter*, Piala Dunia U-20

## **ABSTRACT**

Utama, Singgih. 2026. *Sentiment Analysis on Twitter After the Cancellation of the FIFA U-20 World Cup Hosted by Indonesia Using the Multinomial Naïve Bayes Method*. Undergraduate Thesis. Undergraduate Program. Informatics Engineering Study Program. University of Muhammadiyah Jember.

**Advisors:** Moh. Dasuki, S.Kom., M.Kom.; Habibatul Azizah Al-Faruq, M.Pd.

*This study aims to analyze public sentiment regarding the cancellation of the FIFA U-20 World Cup hosted by Indonesia using the Multinomial Naïve Bayes method. A total of 921 Twitter comments were collected through a crawling process and classified into three sentiment categories: positive, negative, and neutral. The initial dataset showed a significant class imbalance, consisting of 859 positive, 55 negative, and 7 neutral comments. Therefore, the Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the dataset into 429 instances for each class. The text data were then processed through several preprocessing stages, including cleansing, case folding, tokenization, stopwords removal, and stemming, followed by weighting using Term Frequency–Inverse Document Frequency (TF-IDF). The evaluation results using K-Fold Cross Validation demonstrated that the application of SMOTE significantly improved the model's performance. The best performance was achieved in K-Fold 9, obtaining an accuracy of 97.20%, precision of 97.31%, and recall of 97.22%. These findings indicate that the Multinomial Naïve Bayes method combined with TF-IDF weighting and SMOTE is effective for sentiment classification on imbalanced datasets.*

**Keywords:** Sentiment Analysis, Multinomial Naïve Bayes, SMOTE, TF-IDF, Twitter, U-20 World Cup