

ABSTRAK

Andi F, Muhammad. 2026. Analisis Sentimen Pada Media Sosial X menggunakan *Multinomial Naïve Bayes* terhadap Kasus Bocornya Pusat data Nasional. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

Pembimbing : (1) Deni Arifianto M.Kom.,; (2) Ginanjar Abdurrahman, S.Si.,M.Pd. Kasus kebocoran data pada Pusat Data Nasional menjadi salah satu isu yang cukup banyak mendapat perhatian publik. Peristiwa ini kembali terjadi pada Juni 2024, yaitu ketika Pusat Data Nasional mengalami insiden kebocoran data dalam skala besar. Insiden tersebut melibatkan berbagai informasi pribadi yang bersifat sensitif, seperti nama lengkap, Nomor Induk Kependudukan (NIK), hingga data keuangan berupa informasi rekening bank. Hal ini memicu berbagai tanggapan dan opini publik yang disampaikan melalui media sosial, khususnya platform X. Media sosial X menjadi salah satu sarana masyarakat dalam menyampaikan pendapat dan kritik, banyaknya opini tersebut dapat menjadi sumber data yang bermanfaat untuk proses analisis sentimen. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat pada media sosial X tentang kasus bocornya pusat data nasional dan mengukur performa dari *Multinomial Naive Bayes*. Tahapan penelitian berupa pengumpulan data dari X, pelabelan data, *Random Oversampling*, pra proses data, pembobotan *TF-IDF*, klasifikasi *Multinomial Naive Bayes*, Validasi *K-Fold*, dan Evaluasi *Confusion Matrix*. Sebanyak 1264 data diklasifikasi oleh *Multinomial Naive Bayes* dengan pengujian *K Fold Cross Validation* yang menghasilkan akurasi tertinggi di pengujian ke 3 pada *6 fold cross validation* yaitu 82%, untuk *Precision* 78%, *Recall* 89%, dan *F1-score* 84% di label negatif. Sedangkan untuk label positif, nilai *Precision* mencapai 88%, *Recall* 75%, dan *F1-score* 81%

Kata Kunci : Klasifikasi Sentimen, Kebocoran Data, *Multinomial Naïve Bayes*

ABSTRACT

Andi F, Muhammad. 2026. Sentiment Analysis on Social Media X regarding the National Data Center Leak Case using Multinomial Naïve Bayes. Final Project. Undergraduate Program. Informatics Engineering Study Program. Muhammadiyah University of Jember.

Mentor : (1) Deni Arifianto M.Kom.,; (2) Ginanjar Abdurrahman, S.Si.,M.Pd.

The data breach case at the National Data Center has become one of the issues that has received considerable public attention. This incident occurred again in June 2024, when the National Data Center experienced a large-scale data leak. The incident involved various sensitive personal information, such as full names, National Identification Numbers (NIK), and financial data including bank account information. This triggered various public responses and opinions expressed through social media, especially platform X. Social media X has become one of the means for the public to express opinions and criticism, and the large number of opinions can serve as a valuable source of data for the sentiment analysis process. This study aims to classify public sentiment on social media X regarding the national data center leak case and to measure the performance of Multinomial Naive Bayes. The research stages include data collection from X, data labeling, Random Oversampling, data preprocessing, TF-IDF weighting, Multinomial Naive Bayes classification, K-Fold Validation, and Confusion Matrix Evaluation. A total of 1,264 data points were classified by Multinomial Naive Bayes using K-Fold Cross Validation testing, which resulted in the highest accuracy during the tests. The results from testing with K=6 fold cross-validation show that the Multinomial Naive Bayes method achieved the highest accuracy on the 3rd test of the 6-fold cross-validation, which was 82%. For the negative label, Precision was 78%, Recall 89%, and F1-score 84%. Meanwhile, for the positive label, Precision reached 88%, Recall 75%, and F1-score 81%.

Keywords: *Sentiment Classification, Data Breach, Multinomial Naive Bayes*