

## Analisis Sentimen Ulasan Aplikasi Gojek Di Google Play Store Menggunakan Multinomial Naïve Bayes

Ingrida Suryani Lundam<sup>1</sup>, J.B.Budi Darmawan<sup>2</sup>

<sup>1,2</sup> Jurusan Informatika, Universitas Sanata Dharma Yogyakarta

Korespondensi : [b.darmawan@usd.ac.id](mailto:b.darmawan@usd.ac.id)

### ABSTRAK

Perkembangan teknologi di Indonesia telah memfasilitasi berbagai aktivitas manusia, termasuk dalam bidang transportasi melalui layanan berbasis online seperti Gojek, sebuah perusahaan decacorn yang menyediakan berbagai layanan transportasi. Gojek telah mengubah cara masyarakat Indonesia bergerak di perkotaan. Pemahaman terhadap sentimen pengguna terhadap aplikasi Gojek, terutama melalui ulasan di Google Play Store, menjadi kunci dalam meningkatkan dan memperbaiki layanan yang ditawarkan. Penelitian ini bertujuan untuk mengevaluasi kinerja analisis sentimen terhadap 4746 ulasan aplikasi Gojek di Google Play Store menggunakan algoritma Naive Bayes. Metode yang digunakan meliputi proses penerjemahan data, pelabelan menggunakan Vader, preprocessing teks, pembobotan kata dengan TF-IDF, penyeimbangan data dengan SMOTE, dan pengujian model menggunakan K-Fold cross validation. Klasifikasi sentimen dilakukan dengan Multinomial Naive Bayes, dan evaluasi model menggunakan confusion matrix. Hasil Penelitian menunjukkan bahwa pengujian dengan K-fold cross validation  $k = 11$  dan  $\alpha = 0.001$  menghasilkan akurasi tertinggi sebesar 92,25%, precision sebesar 92,16%, recall sebesar 89,71%, dan F1-Score sebesar 90,42%.

**Kata kunci:** Ulasan Aplikasi Gojek, Analisis Sentimen, Naïve Bayes

### ABSTRACT

*The development of technology in Indonesia has facilitated various human activities, including in the field of transportation through online-based services such as Gojek, a decacorn company that provides various transportation services. Gojek has changed the way Indonesian people move around in urban areas. Understanding user sentiment towards the Gojek application, especially through reviews on the Google Play Store, is key to improving and improving the services offered. This study aims to evaluate the performance of sentiment analysis on 4746 Gojek application reviews on the Google Play Store using the Naive Bayes algorithm. The methods used include data translation processes, labeling using Vader, text preprocessing, word weighting with TF-IDF, data balancing with SMOTE, and model testing using K-fold cross validation. Sentiment classification is done using Multinomial Naive Bayes, and model evaluation using a confusion matrix. The research results show that testing with K-fold cross validation  $k = 11$  and  $\alpha = 0.001$  produces the highest accuracy of 92.25%, precision of 92.16%, recall of 89.71%, and F1-score of 90.42%.*

**Keyword :** Gojek Application Review, Sentiment Analysis, Naïve Bayes

### PENDAHULUAN

Perkembangan teknologi di Indonesia telah mendorong berbagai aktivitas manusia untuk beralih ke platform digital, salah satunya dalam bidang transportasi. Munculnya berbagai model transportasi berbasis online di kota-kota besar di Indonesia menunjukkan pergeseran signifikan dalam cara masyarakat bergerak. Gojek, yang didirikan tahun 2010 oleh Nadiem Makarim, adalah salah satu penyedia layanan transportasi online yang paling populer di Indonesia. PT Aplikasi Karya Anak Bangsa (Gojek) telah mengubah cara masyarakat bertransportasi dengan mengintegrasikan teknologi digital dalam layanannya. Selain pandemi Covid-19, permintaan terhadap layanan Gojek meningkat pesat, karena masyarakat beralih dari transportasi umum ke opsi yang lebih aman dan nyaman. Namun, dengan meningkatnya jumlah aplikasi transportasi online seperti Grab, Maxim, dan Blue Bird, persaingan di pasar transportasi digital semakin ketat. Untuk tetap kompetitif, Gojek harus terus berinovasi, meningkatkan layanan, serta merespon *feedback* pengguna dengan cepat dan efektif.

Tanggapan masyarakat terhadap Gojek sangat bervariasi. Beberapa pengguna memuji kemudahan aplikasi, keberagaman layanan, dan kecepatan layanan, sementara yang lain mengkritik keterlambatan pengemudi, kualitas pelayanan yang tidak konsisten, dan harga yang dianggap tinggi. Calon pengguna sering

membandingkan Gojek dengan kompetitor berdasarkan biaya, kenyamanan, dan keamanan. Untuk menjaga loyalitas pengguna dan menarik pengguna baru, Gojek perlu melakukan perbaikan yang berkelanjutan. Analisis sentimen adalah teknik dalam text mining yang digunakan untuk menentukan polaritas sentimen, yaitu positif, netral, dan negatif dari teks. Dalam konteks aplikasi transportasi seperti Gojek, pemahaman terhadap sentimen pengguna melalui ulasan di Google Play Store dapat memberikan wawasan berharga untuk meningkatkan kualitas layanan. Banyak penelitian menunjukkan bahwa algoritma Naïve Bayes, khususnya Multinomial Naïve Bayes, efektif dalam analisis teks dan sentimen [1]. Algoritma Multinomial Naïve Bayes digunakan dalam penelitian ini untuk menganalisis ulasan pengguna Gojek dan memberikan informasi yang lebih baik tentang pengalaman pengguna [2].

**METODE PENELITIAN**

Penelitian ini dilakukan melalui beberapa tahap, mulai dari pengumpulan data, preprocessing, pembobotan TF-IDF, hingga klasifikasi dengan Multinomial Naïve Bayes.

**Pengumpulan Data**

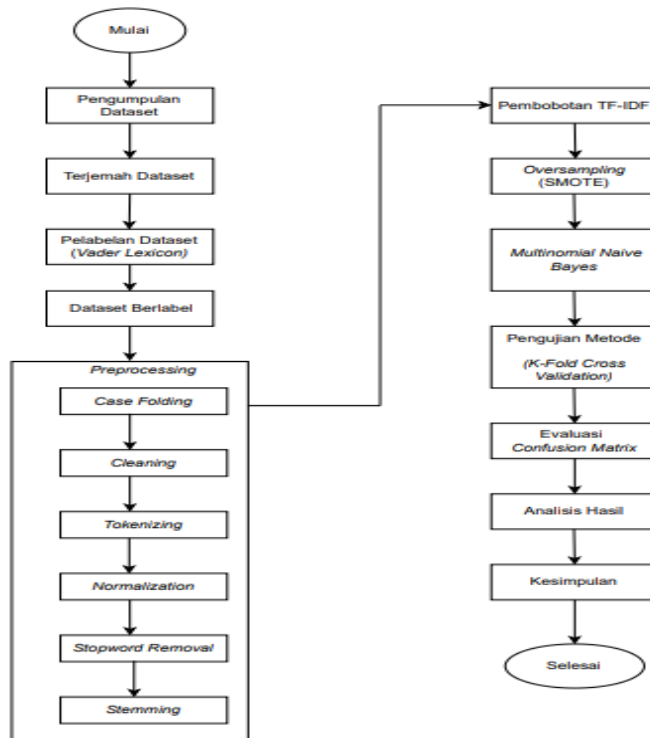
Data yang digunakan adalah ulasan aplikasi Gojek di Google Play Store yang diambil melalui teknik *web scraping*. Sebanyak 6030 ulasan dalam bahasa Indonesia yang diperoleh meliputi informasi seperti username, *score*, *At*, dan *content*. Pada Gambar 1 disajikan contoh dataset hasil *web scraping*.

	userName	score	at	content
0	Michael Lionel Aquino Motombri	5	2024-04-28 05:10:06	Keren
1	Amelia Hussana	1	2024-04-28 05:02:47	akun saya pernah hilang saldo sekitar 3 thn ya...

Gambar 1. Contoh dataset hasil *web scraping*

**Gambaran Umum Penelitian**

Bagian ini menjelaskan tahapan umum penelitian terkait analisis sentimen ulasan aplikasi Gojek di Google Play Store menggunakan algoritma Multinomial Naïve Bayes yang disajikan pada Gambar 2.



Gambar 2. Tahapan penelitian

### Terjemah Data

Setelah dataset berhasil diperoleh, langkah selanjutnya adalah menerjemahkan dataset dari bahasa Indonesia menjadi bahasa Inggris menggunakan *library deep translator* yang mendukung Google Translator API. Contoh data hasil terjemahan disajikan pada Tabel 1.

**Tabel 1.** Data hasil terjemahan

Sebelum terjemahan	Setelah terjemahan
nunggu go car lama banget buat apa kalo murah tapi lama banyak driver yang dekat tapi ga dapet	What's the point of waiting so long for a go car if it's cheap but takes a long time, lots of drivers are close by but don't get one

### Pelabelan Dataset

Pada tahap ini akan dilakukan pelabelan data pada dataset yang sudah dikumpulkan dan menghasilkan sebanyak 4746 ulasan yang berlabel. Proses pelabelan data dilakukan dengan pelabelan Vader Lexico [3]. Pelabelan Vader Lexicon menggunakan data yang sudah melalui tahap terjemahan data. Kemudian data tersebut akan diberi label menjadi 3 kategori yaitu positif, negatif dan netral. Kemudian kategori positif berisi kata yang bermakna positif, dukungan, persetujuan. Untuk kategori negatif berisi kata yang bermakna negatif, ejekan, penolakan. Dan kategori netral berisi kata yang menyampaikan pendapat atau informasi. Pelabelan dataset menghasilkan 2708 label positif, 1037 label negatif dan 1001 label netral.

### Preprocessing

*Preprocessing* adalah langkah awal dalam memproses teks, bertujuan mengubah dokumen menjadi data terstruktur agar dapat diolah oleh *text mining*. Proses ini juga bertujuan meningkatkan akurasi klasifikasi data [4]. Tahapan *preprocessing* yang dilakukan adalah *case folding*, *cleaning*, *tokenizing*, *normalization*, *stopword removal* dan *stemming*.

### Case Folding

*Case folding* untuk merubah semua huruf menjadi huruf kecil, contoh *case folding* disajikan pada Tabel 2.

**Tabel 2.** Hasil tahap *case folding*

Sebelum <i>case folding</i>	Setelah <i>case folding</i>
NUNGGU GO CAR BGT KALO MURAH? DRIVER YG DEKAT TAPI GA DAPETA <sup>2</sup>	nunggu go car bgt kalo murah ? driver yg dekat tapi ga dapeta <sup>2</sup>

### Cleaning

Tanda baca, angka, dan karakter khusus dihapus untuk membersihkan dataset. Contoh hasil *cleaning* disajikan pada Tabel 3.

**Tabel 3.** Hasil tahap *cleaning*

Sebelum <i>cleaning</i>	Setelah <i>cleaning</i>
nunggu go car bgt kalo murah ? driver yg dekat tapi ga dapeta <sup>2</sup>	nunggu go car bgt kalo murah driver yg dekat tapi ga dapeta

### Tokenizing

Proses dimana teks dipecah menjadi token atau kata-kata individual. Contoh *tokenizing* disajikan pada Tabel 4.

Tabel 4. Hasil tahap *tokenizing*

Sebelum <i>tokenizing</i>	Setelah <i>tokenizing</i>
nunggu go car bgt kalo murah driver yg dekat tapi ga dapeta	['nunggu', 'go', 'car', 'bgt', 'kalo', 'murah', 'driver', 'yg', 'ga', 'dapeta']

**Normalization**

Proses dimana mengubah kata tidak baku menjadi kata baku. Contoh hasil *normalization* disajikan pada Tabel 5.

Tabel 5. Hasil tahap *normalization*

Sebelum <i>normalization</i>	Setelah <i>normalization</i>
['nunggu', 'go', 'car', 'bgt', 'kalo', 'murah', 'driver', 'yg', 'ga', 'dapeta']	['nunggu', ' <b>pergi</b> ', 'car', ' <b>banget</b> ', ' <b>kalau</b> ', 'murah', 'driver', ' <b>yang</b> ', ' <b>tidak</b> ', ' <b>dapat</b> ']

**Stopword Removal**

Proses menghapus kata-kata umum yang tidak memiliki nilai informasi tinggi. Contoh hasil *stopword removal* disajikan pada Tabel 6.

Tabel 6. Hasil tahap *stopword removal*

Sebelum <i>stopword removal</i>	Setelah <i>stopword removal</i>
['nunggu', 'pergi', 'car', 'banget', 'kalau', 'murah', 'driver', 'yang', 'tidak', 'dapat']	nunggu <b>pergi</b> car <b>kalau</b> murah driver <b>yang tidak dapat</b>

**Stemming**

Proses mengubah kata-kata menjadi bentuk dasar (root word). Contoh hasil *stemming* disajikan pada Tabel 7.

Tabel 7. Hasil tahap *stemming*

Sebelum <i>stemming</i>	Setelah <i>stemming</i>
nunggu pergi car kalau murah driver yang tidak dapat dan	nunggu pergi car kalau murah driver yang tidak dapat

**Pembobotan TF-IDF**

Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode untuk menghitung bobot setiap kata. TF-IDF merupakan gabungan dari Term Frequency (TF) dan Inverse Dokumen Frequency (IDF). TF adalah frekuensi kemunculan kata dalam dokumen, semakin tinggi TF maka semakin besar bobot dokumen [5]. Rumus TF-IDF disajikan pada persamaan (1).

$$w_{i,j} = tf_{i,j} \times idf_i \quad (1)$$

Keterangan :

$w_{i,j}$  : Bobot kata-*i* pada dokumen ke-*j*

$tf_{i,j}$  : Jumlah kata-*i* pada dokumen ke-*j*

$idf_i$  : Inverse document frequency dari kata-*i*

Untuk mengimplementasikan rumus TF-IDF, maka diperlukan rumus untuk menghitung IDF terlebih dahulu [6]. Berikut rumus persamaan untuk menghitung IDF yang ditunjukkan pada persamaan (2).

$$idf_i = \log \left( \frac{N+1}{df_{i+1}} \right) + 1 \quad (2)$$

Keterangan :

$idf_i$  : Inverse document frequency dari kata  $i$

$N$  : Jumlah keseluruhan dokumen

$df_i$  : Jumlah dokumen yang mengandung kata ke- $i$

### SMOTE

Metode *Synthetic Minority Oversampling Method* (SMOTE) digunakan untuk menangani dataset yang tidak seimbang dengan menambah data sintetis pada kelas minoritas agar seimbang dengan kelas mayoritas [7]. Metode ini diterapkan karena ketidakseimbangan antara jumlah data sentimen positif, netral, dan negatif. Oversampling dilakukan untuk menyeimbangkan jumlah data dengan menambahkan kelas yang lebih kecil. Proses penyeimbangan menghasilkan jumlah dataset dengan jumlah 2708 untuk masing-masing label positif, negatif dan netral.

### Algoritma Multinomial Naïve Bayes

*Multinomial Naïve Bayes* adalah metode klasifikasi yang menggunakan probabilitas yang sangat efektif untuk data teks dan frekuensi kejadian [8]. Metode ini mengasumsikan independensi kuat antara fitur dan bekerja berdasarkan teorema *Bayes*, dengan tahapan menghitung jumlah kelas atau label, menghitung jumlah kelas yang sesuai dengan kasus yang diamati, mengalikan variabel-variabel kelas, dan membandingkan hasilnya untuk menentukan kelas yang paling mungkin [9]. Berikut rumus untuk menghitung *Multinomial Naïve Bayes* yang ditunjukkan pada persamaan (3).

$$P(H|X) = \frac{P(X|H)}{P(X)} P(H) \quad (3)$$

Keterangan:

$X$  : Data dengan kelas yang belum diketahui

$H$  : Hipotesis data merupakan suatu kelas spesifik

$P(H|X)$  : Probabilitas hipotesis  $H$  berdasarkan kondisi  $X$  (Posterior Probabilitas)

$P(H)$  : Probabilitas hipotesis  $H$  (Prior probabilitas)

$P(X|H)$  : Probabilitas  $X$  berdasarkan kondisi pada hipotesis

$P(X)$  : Probabilitas  $X$ .

Posterior dibandingkan antara kelas untuk menentukan klasifikasi. Rumus probabilitas posterior dalam *Multinomial Naïve Bayes* ditunjukkan pada persamaan (4).

$$P(C|d) = \frac{P(C) \cdot P(d|C)}{P(d)} \quad (4)$$

Keterangan :

$P(C|d)$  : Probabilitas sampel  $d$  dalam kelas  $C$

$P(C)$  : Probabilitas prior kelas  $C$

$P(d|C)$  : Likelihood, probabilitas sampel  $d$  dalam kelas  $C$

$P(d)$  : Evidence, probabilitas keseluruhan dari sampel  $d$  muncul diseluruh kelas

Untuk perhitungan likelihood dihitung menggunakan persamaan (5).

$$P(d|C) = \prod_{i=1}^n p(x_i|C) f_i \quad (5)$$

Keterangan :

$N$  : Jumlah fitur dalam sampel  $d$ .



- $x_i$  : kata ke- $i$  dalam sampel  $d$ .  
 $f_i$  : Jumlah kemunculan fitur  $x_i$  dalam sampel  $d$ .  
 $P(x_i | C)$  : probabilitas fitur  $x_i$  muncul dalam kelas  $C$ .

### Evaluasi Model

Model dievaluasi menggunakan *confusion matrix* dan kinerjanya diukur berdasarkan akurasi, precision, recall, dan F1-score [10].

#### Accuracy

*Accuracy* memberikan gambaran seberapa baik model dapat mengklasifikasi data dengan benar. Rumus untuk menghitung *accuracy* ditunjukkan pada persamaan (6).

$$accuracy = \frac{True\ Positive + False\ Negatif}{True\ Positive + False\ Positif + True\ Negatif + False\ Negatif} \times 100\ \% \quad (6)$$

#### Precision

*Precision* mencerminkan sejauh mana model memberikan prediksi yang akurat ketika mengklasifikasi data sebagai positif. Rumus untuk menghitung *precision* ditunjukkan pada persamaan (7).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positif} \times 100\ \% \quad (7)$$

#### Recall

*Recall* menggambarkan seberapa baik model mengenali atau menentukan data positif dalam dataset. Rumus untuk menghitung *recall* ditunjukkan pada persamaan (8).

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negatif} \quad (8)$$

#### F1-score

*F1-score* adalah sebuah matrik yang menggunakan nilai precision dan recall untuk memberikan gambaran keseluruhan tentang kinerja suatu model atau sistem. Rumus untuk menghitung *F1-score* ditunjukkan pada persamaan (9).

$$F1 - score = 2 \times \frac{Precision \times recall}{Precision + recall} \quad (9)$$

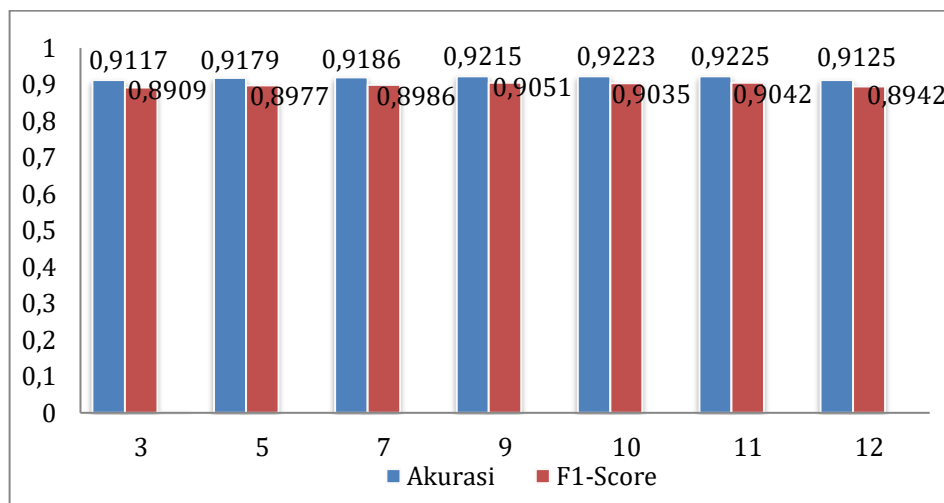
## HASIL DAN ANALISIS

Skenario pengujian dalam pengujian ini menggunakan variasi K-fold dengan nilai  $k = 3, 5, 7, 9, 10, 11, 12$  dan variasi nilai  $alpha = 0,001, 0,01, 0,1, 1, 10$ . Dari hasil pengujian dengan skenario ini untuk semua variasi nilai  $k$  pada K-fold, hasil akurasi optimal diperoleh saat nilai  $alpha = 0,001$ . Pada Tabel 8 dan Gambar 3 disajikan hasil akurasi optimal beserta F1-score untuk setiap variasi nilai  $k$  pada K-fold dengan nilai  $alpha = 0,001$ .

**Tabel 8.** Hasil pengujian untuk setiap variasi nilai  $k$  pada K-fold dengan nilai  $alpha = 0,001$ .

K-Fold	Alpha	Akurasi	Precision	Recall	F1-Score
3	0,001	0,9117	0,9112	0,8836	0,8909
5	0,001	0,9179	0,9177	0,8900	0,8977
7	0,001	0,9186	0,9185	0,8903	0,8986
9	0,001	0,9215	0,9212	0,8985	0,9051
10	0,001	0,9223	0,9215	0,8958	0,9035

11	0,001	0,9225	0,9216	0,8971	0,9042
12	0,001	0,9125	0,9116	0,8871	0,8942



**Gambar 3.** Hasil akurasi optimal beserta F1-score untuk setiap variasi nilai k pada K-fold dengan nilai alpha = 0,001.

Dari hasil pengujian yang disajikan pada Tabel 8. dan Gambar 3. di atas menunjukkan model dengan nilai akurasi yang optimal diperoleh saat K-fold dengan nilai  $k = 11$  dan nilai alpha 0,001 dengan akurasi sebesar 92,25% dan *F1-Score* sebesar 90,42%. Dari grafik pada Gambar 3. terlihat bahwa peningkatan nilai k dari = 11 menjadi nilai k = 12 mulai menunjukkan penurunan nilai akurasi.

## KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, analisis sentimen ulasan aplikasi Gojek di Google Play Store menggunakan algoritma *Multinomial Naïve Bayes* dengan total 4746 data ulasan menghasilkan model yang optimal menggunakan metode K-Fold *cross validation* dengan  $k = 11$  dan nilai alpha 0,001. Model ini mampu mencapai akurasi sebesar 92,25% dengan *precision* sebesar 92,16%, *recall* sebesar 89,71%, dan *F1-Score* sebesar 90,42%. Hasil pemodelan ini diharapkan dapat membantu memprediksi pelabelan ulasan sebagai informasi berharga bagi manajemen untuk meningkatkan kualitas layanan Gojek, meningkatkan kepuasan pengguna, dan menarik lebih banyak pengguna baru.

## DAFTAR PUSTAKA

- [1] W. Zhang and F. Gao, "An improvement to naive bayes for text classification," *Procedia Engineering*, vol. 15, pp. 2160–2164, 2011.
- [2] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5432–5435, 2009.
- [3] Y. Asri, W. N. Suliyanti, D. Kuswardani, and M. Fajri, "Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile." Petir, 2022.
- [4] N. Agustina, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store: The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews On Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 1, pp. 47–54, 2022.
- [5] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, 2021.
- [6] S. Jansen, *Hands-On Machine Learning for Algorithmic Trading: Design and implement investment strategies based on smart algorithms that learn from data using Python*. Packt Publishing Ltd, 2018.
- [7] R. Siringoringo and I. K. Jaya, "Ensemble Learning Dengan Metode Smotebagging Pada Klasifikasi Data Tidak Seimbang," *Journal Information System Development (ISD)*, vol. 3, no. 2, 2018.

- 
- [8] F. Nurhuda, S. W. Sihwi, and A. Doewes, “Analisis sentimen masyarakat terhadap calon Presiden Indonesia 2014 berdasarkan opini dari Twitter menggunakan metode Naive Bayes Classifier,” *ITSmart: Jurnal Teknologi dan Informasi*, vol. 2, no. 2, pp. 35–42, 2016.
  - [9] A. P. Wibawa, M. Guntur, A. Purnama, M. F. Akbar, and F. A. Dwiyanto, “Metode-metode klasifikasi,” in *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, 2018.
  - [10] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, “Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naive Bayes,” *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, pp. 422–430, 2021.