

**ANALISIS SENTIMEN TERHADAP REVIEW APLIKASI FLIP
MENGUNAKAN METODE NAIVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NURAINI YULIANITA

21.11.4199

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN TERHADAP REVIEW APLIKASI FLIP
MENGUNAKAN METODE NAIVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NURAINI YULIANITA

21.11.4199

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP REVIEW APLIKASI FLIP
MENGUNAKAN METODE NAIVE BAYES**

yang disusun dan diajukan oleh

Nuraini Yulianita

21.11.4199

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Februari 2025

Dosen Pembimbing,



I Made Artha Agastya, S.T., M.Eng., Ph.D.

NIK. 190302352

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN TERHADAP REVIEW APLIKASI FLIP
MENGGUNAKAN METODE NAIVE BAYES

yang disusun dan diajukan oleh

Nuraini Yulianita

21.11.4199

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Rizqi Sukma Kharisma, S.Kom., M.Kom.
NIK. 190302215

Melany Mustika Dewi, S.Kom., M.Kom.
NIK. 190302455

I Made Artha Agastya, S.T., M.Eng., Ph.D.
NIK. 190302352



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nuraini Yulianita
NIM : 21.11.4199

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Terhadap Review Aplikasi Flip Menggunakan Metode Naïve Bayes

Dosen Pembimbing : I Made Artha Agastya, S.T, M.Eng, Ph.D.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 Februari 2025

Yang Menyatakan,



Nuraini Yulianita

KATA PENGANTAR

Puji syukur saya ucapkan kepada Allah SWT karena atas nikmat dan rahmat Nya, penulis dapat menyelesaikan skripsi ini sebagai syarat untuk mencapai gelar Sarjana Komputer Program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta. Selama proses penyusunan skripsi ini, tidak terlepas dari beberapa bantuan, saran, kritik dan dukungan.

Dalam kesempatan ini, penulis berterima kasih kepada:

1. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
2. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Program Studi Informatika dan Bapak Arif Akbarul Huda, S.Si., M.Eng., selaku Sekretaris Program Studi Informatika.
3. Bapak I Made Artha Agastya, S.T, M.Eng, Ph.D selaku Dosen Pembimbing yang telah memberikan arahan, motivasi dan saran serta telah bersedia meluangkan waktu untuk membimbing penulis hingga akhir penulisan skripsi ini dengan baik.
4. Seluruh civitas akademik, khususnya dosen dan staf karyawan Fakultas Ilmu Komputer telah memberikan bantuan dan ilmu selama penulis melakukan perkuliahan.
5. Kedua orang tua tercinta dan keluarga yang selalu memanjatkan doa, dukungan dan memberi motivasi untuk penulis selama penyusunan skripsi.
6. Teman-teman Informatika Angkatan 2021 yang telah berjuang Bersama untuk menyelesaikan skripsi.

Yogyakarta, 8 Maret 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH	xii
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	10
2.2.1 Analisis Sentimen	10

2.2.2	<i>Natural Language Processing</i>	10
2.2.3	<i>Aplikasi Fintech</i>	10
2.2.4	<i>Term Frequency-Inverse Document Frequency</i>	10
2.2.5	Naive Bayes	11
2.2.6	<i>Oversampling & Undersampling</i>	11
2.2.7	Evaluasi Kinerja Model	16
BAB III METODE PENELITIAN		17
3.1	Objek Penelitian	17
3.2	Alur Penelitian	17
3.2.1	Pengumpulan Data	18
3.2.2	<i>Preprocessing Data</i>	19
3.2.3	Visualisasi Data	20
3.2.4	Pembagian Dataset	20
3.2.5	Transformasi Data (TF-IDF)	20
3.2.6	Klasifikasi	21
3.2.7	Penanganan Ketidakseimbangan Data	21
3.2.7	Evaluasi	22
BAB IV HASIL DAN PEMBAHASAN		23
4.1	<i>Sample</i>	23
4.1.1	<i>Scraping Data</i>	23
4.2	<i>Explore</i>	23
4.3	<i>Modify</i>	24
4.3.1	<i>Labeling</i>	24
4.3.2	<i>Case Folding</i>	26
4.3.3	<i>Stopword Removal</i>	27

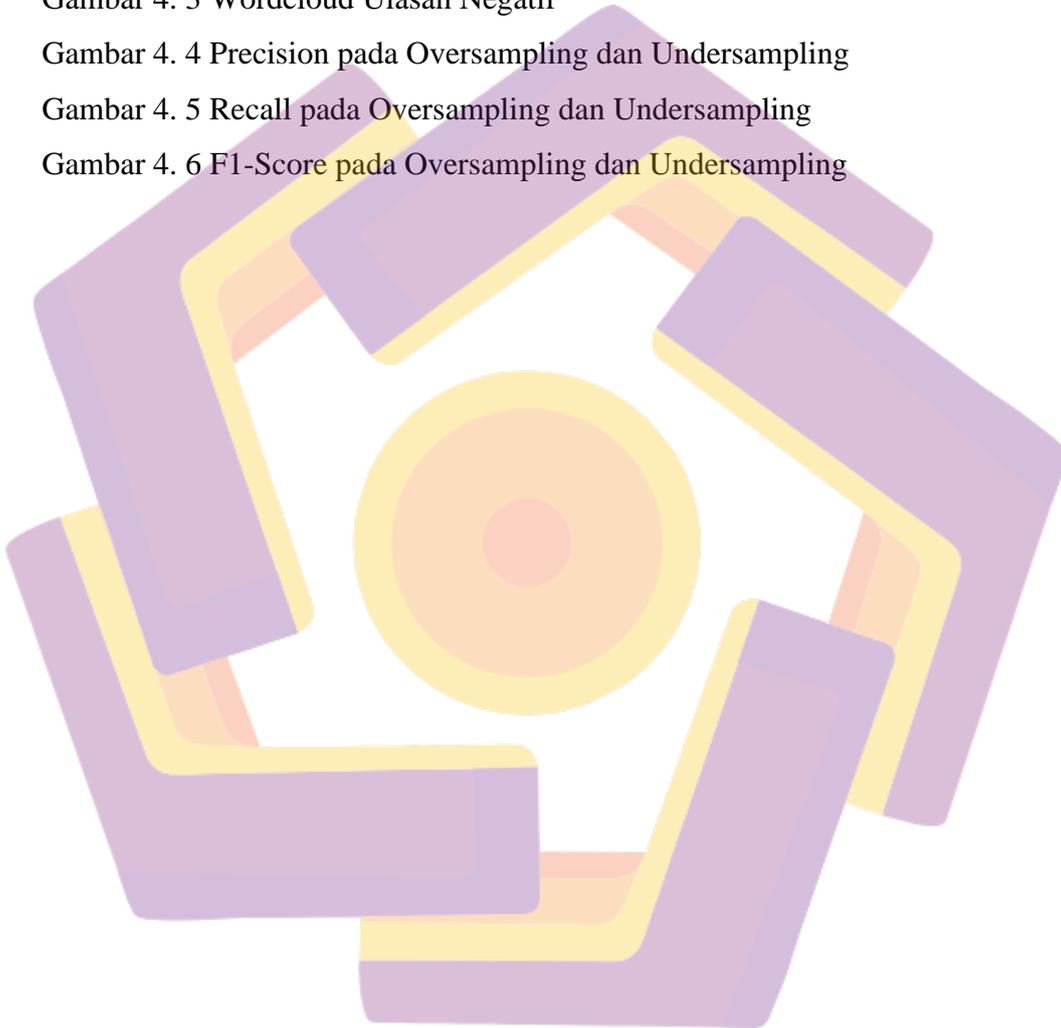
4.3.4 <i>Tokenizing</i>	28
4.3.5 <i>Stemming</i>	28
4.3.6 <i>Wordcloud</i>	28
4.4 Model	30
4.4.1 Pembagian Dataset	30
4.4.2 TF-IDF	30
4.4.3 Klasifikasi	33
4.4.4 Evaluasi	34
BAB V PENUTUP	38
5.1 Kesimpulan	38
5.2 Saran	38
REFERENSI	39

DAFTAR TABEL

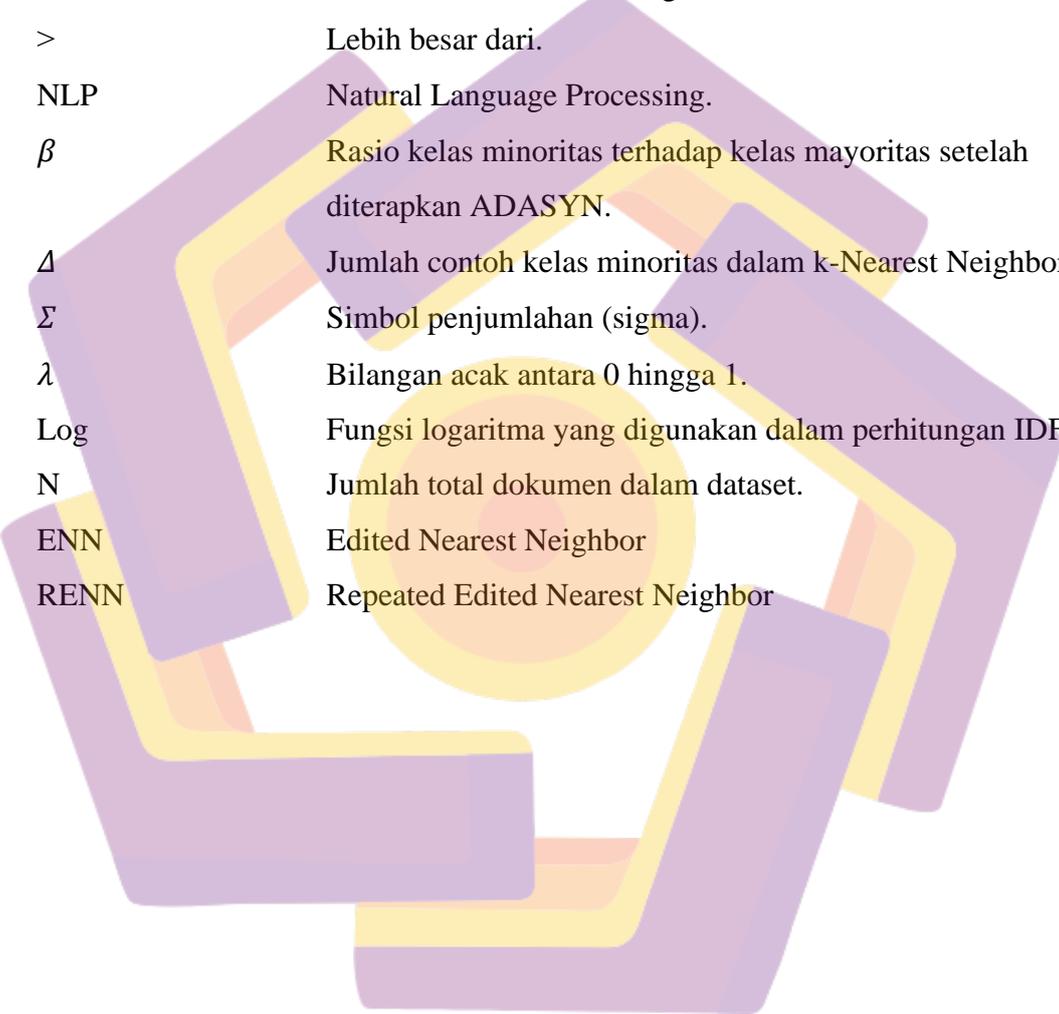
Tabel 2. 1 Keaslian Penelitian	7
Tabel 4. 1 Hasil Scraping Ulasan dengan Rating	24
Tabel 4. 2 Hasil Labeling	25
Tabel 4. 3 Hasil Case Folding	27
Tabel 4. 4 Hasil Stopword Removal	27
Tabel 4. 5 Hasil Tokenizing	28
Tabel 4. 6 Hasil Pembagian Dataset	30
Tabel 4. 7 Contoh Data Ulasan	31
Tabel 4. 8 Tabel Perhitungan TF	31
Tabel 4. 9 Tabel Perhitungan IDF	32
Tabel 4. 10 Tabel Perhitungan TF-IDF	33
Tabel 4. 11 Hasil Klasifikasi Algoritma Naïve Bayes	33

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	18
Gambar 4. 1 Frekuensi Ulasan Berdasarkan Sentimen	26
Gambar 4. 2 Wordcloud Ulasan Positif	29
Gambar 4. 3 Wordcloud Ulasan Negatif	29
Gambar 4. 4 Precision pada Oversampling dan Undersampling	35
Gambar 4. 5 Recall pada Oversampling dan Undersampling	36
Gambar 4. 6 F1-Score pada Oversampling dan Undersampling	37

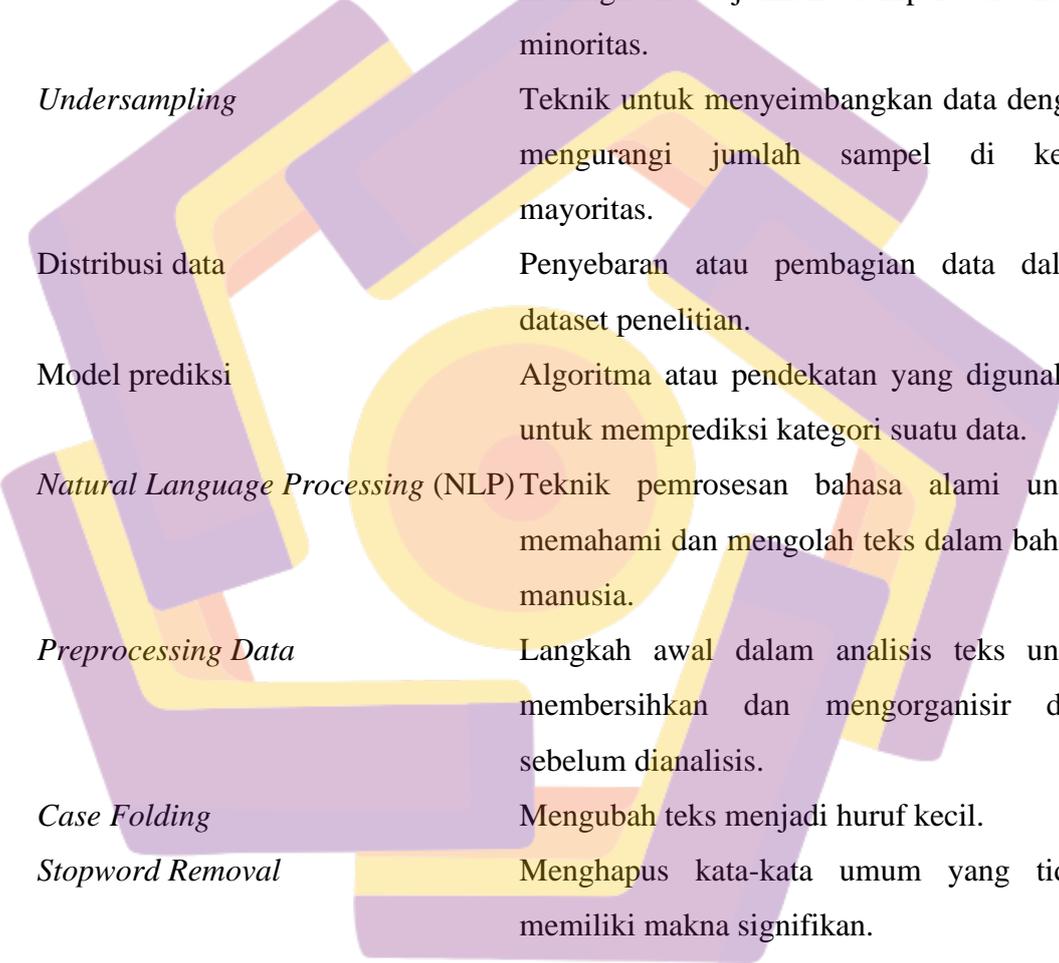


DAFTAR LAMBANG DAN SINGKATAN



TF-IDF	Term Frequency-Inverse Document Frequency.
ADASYN	Adaptive Synthetic Sampling Approach
K-Means SMOTE	K-Means Synthetic Minority Oversampling Technique.
\leq	Lebih kecil atau sama dengan.
$>$	Lebih besar dari.
NLP	Natural Language Processing.
β	Rasio kelas minoritas terhadap kelas mayoritas setelah diterapkan ADASYN.
Δ	Jumlah contoh kelas minoritas dalam k-Nearest Neighbors.
Σ	Simbol penjumlahan (sigma).
λ	Bilangan acak antara 0 hingga 1.
Log	Fungsi logaritma yang digunakan dalam perhitungan IDF.
N	Jumlah total dokumen dalam dataset.
ENN	Edited Nearest Neighbor
RENN	Repeated Edited Nearest Neighbor

DAFTAR ISTILAH



Analisis sentimen	Teknik untuk mengidentifikasi dan mengekstrak informasi tentang sikap, pandangan, dan emosi dari teks
<i>Oversampling</i>	Teknik untuk menyeimbangkan data dengan meningkatkan jumlah sampel di kelas minoritas.
<i>Undersampling</i>	Teknik untuk menyeimbangkan data dengan mengurangi jumlah sampel di kelas mayoritas.
Distribusi data	Penyebaran atau pembagian data dalam dataset penelitian.
Model prediksi	Algoritma atau pendekatan yang digunakan untuk memprediksi kategori suatu data.
<i>Natural Language Processing (NLP)</i>	Teknik pemrosesan bahasa alami untuk memahami dan mengolah teks dalam bahasa manusia.
<i>Preprocessing Data</i>	Langkah awal dalam analisis teks untuk membersihkan dan mengorganisir data sebelum dianalisis.
<i>Case Folding</i>	Mengubah teks menjadi huruf kecil.
<i>Stopword Removal</i>	Menghapus kata-kata umum yang tidak memiliki makna signifikan.
<i>Tokenizing</i>	Memecah teks menjadi kata atau frasa.
<i>Stemming</i>	Mengubah kata menjadi bentuk dasarnya.
Aplikasi Flip	Aplikasi keuangan berbasis Android untuk transfer antarbank tanpa biaya administrasi.
TF-IDF	Teknik untuk mengubah teks menjadi bentuk numerik berdasarkan pentingnya kata dalam

	suatu dokumen dibandingkan dengan dataset keseluruhan.
Naive Bayes	Algoritma berbasis probabilitas untuk klasifikasi teks yang menggunakan Teorema Bayes.
ADASYN	Teknik oversampling yang berfokus pada daerah dengan kepadatan data minoritas rendah.
SMOTE	Metode untuk menghasilkan data sintetis berdasarkan tetangga terdekat dari kelas minoritas.
<i>Random Over Sampler</i>	Teknik sederhana yang mereplikasi data minoritas secara acak.
<i>Random Under Sampler</i>	Teknik yang menghapus data dari kelas mayoritas secara acak.
Tomek Links	Teknik yang menghapus sampel mayoritas yang berpasangan dengan sampel minoritas di dekat garis batas antar kelas.
<i>Repeated Nearest Neighbour</i>	Teknik yang menghapus sampel mayoritas yang salah diklasifikasikan berdasarkan tetangga terdekatnya.
<i>Feedback</i>	Umpan balik atau ulasan dari pengguna.
<i>Scraping Data</i>	Teknik pengumpulan data secara otomatis dari suatu sumber online.
<i>Labeling</i>	Proses memberi label atau kategori pada data.
<i>Wordcloud</i>	Representasi visual dari kata-kata yang sering muncul.
Transformasi data	Proses mengubah data menjadi format yang lebih dapat diproses.
Bobot kata	Tingkat kepentingan suatu kata dalam teks.

Klasifikasi	Proses mengelompokkan data ke dalam kategori tertentu.
Klaster	Pengelompokan data berdasarkan kesamaan tertentu.
Dataset	Kumpulan data yang digunakan dalam penelitian.
Pelatihan model	Proses melatih model dengan data yang tersedia.
Evaluasi model	Proses mengukur kinerja model yang telah dilatih.
<i>Precision</i>	Ukuran keakuratan prediksi positif dari model.
<i>Recall</i>	Ukuran kemampuan model dalam menangkap semua data positif.
<i>F1-Score</i>	Ukuran keseimbangan antara precision dan recall dalam model klasifikasi.
<i>Google Play Store</i>	Platform distribusi digital yang digunakan untuk mengunduh aplikasi.
Metrik Euclidean	Suatu ukuran jarak antara dua titik dalam ruang Euclidean

INTISARI

Dalam era digital, ulasan pengguna memainkan peran penting dalam menentukan kesuksesan aplikasi. Flip, sebagai salah satu dompet digital di Indonesia, menerima berbagai ulasan dari penggunanya yang mencerminkan pengalamannya. Namun, ketidakseimbangan data antara ulasan positif dan negatif menjadi tantangan dalam analisis sentimen, yang dapat memengaruhi akurasi klasifikasi. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Flip menggunakan metode Naive Bayes. Untuk mengatasi ketidakseimbangan data, teknik *oversampling* dan *undersampling* diterapkan, sementara *preprocessing* seperti *tokenizing*, *stemming*, dan *stopword removal* digunakan guna meningkatkan akurasi model. Data ulasan diperoleh dari *Google Play Store* dan diklasifikasikan ke dalam sentimen positif dan negatif. Hasil penelitian menunjukkan bahwa model Naive Bayes dengan ADASYN memberikan performa terbaik dengan *precision* 0,89, *recall* 0,88, dan *F1-score* 0,88, yang membuktikan bahwa teknik ini efektif dalam menangani ketidakseimbangan data. Penelitian ini memberikan wawasan bagi pengembang aplikasi Flip dalam memahami sentimen pengguna serta meningkatkan layanan berdasarkan masukan pengguna. Selain itu, hasil penelitian ini dapat menjadi referensi bagi studi analisis sentimen lainnya, khususnya dalam pengolahan data teks dan pengembangan model klasifikasi sentimen. Untuk penelitian selanjutnya, disarankan untuk menguji metode klasifikasi lain dan memperluas cakupan dataset guna meningkatkan generalisasi model.

Kata kunci: Sentimen, Naive Bayes, ketidakseimbangan data, ADASYN, klasifikasi.

ABSTRACT

In the digital era, user reviews play an important role in determining the success of an app. Flip, as one of the digital wallets in Indonesia, receives various reviews from its users that reflect their experience. However, data imbalance between positive and negative reviews is a challenge in sentiment analysis, which can affect classification accuracy. This research aims to analyze the sentiment of Flip app user reviews using the Naive Bayes method. To overcome data imbalance, oversampling and undersampling techniques are applied, while preprocessing such as tokenization, stemming, and stopword removal are used to improve model accuracy. Review data was obtained from Google Play Store and classified into positive and negative sentiments. The results show that the Naive Bayes model with ADASYN provides the best performance with a precision of 0.89, recall of 0.88, and F1-score of 0.88, which proves that this technique is effective in handling data imbalance. This research provides insights for Flip app developers in understanding user sentiment and improving services based on user feedback. In addition, the results of this research can be a reference for other sentiment analysis studies, especially in text data processing and sentiment classification model development. For future research, it is recommended to test other classification methods and expand the scope of the dataset to improve the generalizability of the model.

Keyword: Sentiment, Naïve Bayes, data imbalance, ADASYN, classification