

**ANALISIS SENTIMEN ULASAN PENGGUNA
APLIKASI KREDIVO MENGGUNAKAN
ALGORITMA NAIVE BAYES DAN LOGISTIC
REGRESSION**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Mencapai Gelar Sarjana Program Studi Informatika
Fakultas Sains, Teknologi, Dan Kesehatan
Universitas Sahid Surakarta



Disusun Oleh:

Diki Dwi Anggara
NIM. 2022061014

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS, TEKNOLOGI, DAN KESEHATAN
UNIVERSITAS SAHID SURAKARTA
2026**

LEMBAR PERSETUJUAN

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI KREDIVO MENGUNAKAN ALGORITMA NAIVE BAYES DAN LOGISTIC REGRESSION

Disusun oleh:

Diki Dwi Anggara

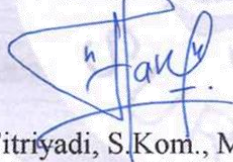
NIM 2022061014

Skripsi ini telah disetujui untuk dipertahankan

Di hadapan dewan penguji

Pada tanggal 13 Agustus 2026

Pembimbing 1



Farid Fitriyadi, S.Kom., M.Kom.

NIDN.0607058902

Pembimbing 2



Ir. Dahlan Susilo, M.Kom.

NIDN.0614016701

Mengetahui,

Ketua Program Studi




Diyah Ruswanti, S.Kom., M.Kom.

NIDN.0027018101

LEMBAR PENGESAHAN

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI KREDIVO MENGGUNAKAN ALGORITMA NAIVE BAYES DAN LOGISTIC REGRESSION

Disusun oleh:

DIKI DWI ANGGARA
NIM.2022061014

Tugas Akhir ini telah diterima dan disahkan
pada hari Kamis tanggal 13 Agustus 2026

Dewan Penguji

Penguji 1

Farid Fitriyadi, S.Kom., M.Kom.
NIDN. 0607058902

Penguji 2

Ir. Dahlan Susilo, M.Kom.
NIDN. 0614016701

Penguji 3

Dr. Sri Huning Anwariningsih, S.T., M.Kom.
NIDN. 0017067901

Mengetahui,

**Kepala Program Studi
Prodi Informatika**

(Diyah Ruswanti, S.Kom., M.Kom.)
NIDN. 0027018101

**Dekan Fakultas Sains,
Teknologi dan Kesehatan**

(apt Ahwan, S.Farm., M.Sc.)
NIDN. 0626088401

SURAT PERNYATAAN ORISINALITAS KARYA ILMIAH

Saya yang bertanda tangan dibawah ini:

Nama : Diki Dwi Anggara

NIM : 2022061014

Program Studi : Informatika

Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa skripsi yang berjudul Analisis Sentimen Ulasan Pengguna Aplikasi Kredivo menggunakan Algoritma Naive Bayes dan Logistic Regression adalah benar karya yang saya susun sendiri. Apabila dikemudian hari saya terbukti melakukan tindakan menyalin atau meniru hasil karya orang lain seolah-olah hasil pemikiran saya sendiri, saya bersedia dan menerima sanksi sesuai peraturan yang berlaku di Universitas Sahid Surakarta termasuk pencabutan gelar yang telah saya peroleh.

Demikian surat pernyataan yang saya buat dengan sebenar-benarnya dan apabila dikemudian hari terbukti melakukan kebohongan maka saya sanggup menanggung segala konsekuensinya.

Surakarta, 18 Juni 2026

Yang membuat pernyataan



Diki Dwi Anggara
NIM.2022061014

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai Civitas Akademis Universitas Sahid Surakarta, saya yang bertanda tangan dibawah ini:

Nama : Diki Dwi Anggara

NIM : 2022061014

Program Studi : Informatika

Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Sahid Surakarta Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalti Free Right*) atas skripsi saya yang berjudul:

"ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI KREDIVO MENGGUNAKAN ALGORITMA NAIVE BAYES DAN LOGISTIC REGRESSION"

Beserta desain atau perangkat (jika ada). Berhak menyimpan, mengalihkan bentuk, mengalihmediakan, mengelola dalam bentuk pangkalan data (*database*), merawat serta mempublikasikan karya ilmiah saya selama tetap mencantumkan nama saya sebagai penulis (*author*) dan pembimbing sebagai *co-author* atau pencipta dan juga sebagai pemilik hak cipta.

Demikian pernyataan yang saya buat dengan sesungguhnya secara sehat dan sadar tanpa ada paksaan dari pihak manapun.

Surakarta, 18 Juni 2026

Yang membuat pernyataan



Diki Dwi Anggara
NIM.2022061014

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan Laporan Skripsi yang berjudul “Analisis Sentimen Ulasan Pengguna Aplikasi Kredivo menggunakan Algoritma Nive Bayes dan Logistic Regression”. Laporan ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Informatika pada Fakultas Sains, Teknologi dan Kesehatan Universitas Sahid Surakarta. Penyusunan proposal ini tidak terlepas dari bantuan, arahan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Ibu Dr. Sri Huning Anwariningsih, S.T., M.Kom., selaku Rektor Universitas Sahid Surakarta.
2. Bapak Apt. Ahwan, S.Farm., M.Sc., selaku Dekan Fakultas Sains, Teknologi dan Kesehatan Universitas Sahid Surakarta.
3. Ibu Diyah Ruswanti, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Sahid Surakarta.
4. Bapak Farid Fitriyadi, S.Kom., M.Kom., selaku Dosen Pembimbing I.
5. Bapak Ir. Dahlan Susilo, M.Kom., selaku Dosen Pembimbing II.
6. Seluruh dosen Program Studi Informatika Universitas Sahid Surakarta yang telah memberikan ilmu, wawasan, serta dukungan selama perkuliahan.
7. Bapak, Ibu, dan seluruh keluarga tercinta yang senantiasa memberikan doa, dukungan moral, dan semangat.
8. Teman-teman serta pihak lain yang tidak dapat penulis sebutkan satu per satu, yang telah membantu dalam proses penyusunan proposal skripsi ini.

Besar harapan penulis laporan ini dapat berguna bagi pembaca dan Universitas Sahid Surakarta.

Surakarta, 2026

LEMBAR PERSEMBAHAN

Dengan penuh rasa syukur ke hadirat Allah SWT atas segala rahmat, nikmat, kesehatan, kekuatan, serta kemudahan yang telah diberikan, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Ulasan Pengguna Aplikasi Kredivo Menggunakan Algoritma Naive Bayes dan Logistic Regression”. Skripsi ini bukan hanya menjadi salah satu syarat untuk menyelesaikan pendidikan, tetapi juga menjadi bukti dari sebuah perjalanan panjang yang di dalamnya terdapat doa, perjuangan, kesabaran, air mata, dukungan, serta banyak orang yang telah memberikan arti dalam setiap prosesnya. Dengan segala kerendahan hati, karya sederhana ini penulis persembahkan kepada:

1. Teruntuk orang tua tercinta, Ibu Warmi dan Bapak Wasino. Terima kasih atas kasih sayang yang tak terbatas, doa yang tak pernah putus, serta perjuangan dan pengorbanan yang luar biasa untuk penulis. Terima kasih telah menjadi tempat pulang ternyaman dan sumber kekuatan terbesar di setiap keadaan. Semoga Ibu dan Bapak selalu diberikan kesehatan, kebahagiaan, umur panjang, dan keberkahan. Karya ini penulis persembahkan sebagai salah satu wujud bakti dan rasa terima kasih penulis.
2. Kepada saudara tercinta, kakak saya Retno Fitriani dan adik saya Tri Prasetya serta Catur Ramadhani. Terima kasih telah memberikan semangat, keceriaan, dan warna dalam hidup penulis. Dukungan dan kebersamaan yang terjalin selama ini menjadi salah satu motivasi bagi penulis untuk terus berusaha dan menyelesaikan studi ini.
3. Kepada sahabat-sahabat seperjuangan, Rifandy, Latif, Ilham, Abidin, dan seluruh angkatan 2022. Terima kasih telah menjadi teman yang saling menguatkan di masa-masa perkuliahan. Terima kasih atas kebersamaan, bantuan, canda tawa, dan cerita yang telah dilalui bersama. Semoga setelah perjalanan ini selesai, kita semua diberikan jalan kesuksesan di bidang masing-masing.

4. Kepada seluruh penghuni kos Liah. Terima kasih telah menjadi lingkungan yang mendukung dan penuh kehangatan selama masa perkuliahan. Terima kasih atas setiap diskusi, bantuan, kebersamaan, dan semangat yang diberikan hingga penulis dapat menyelesaikan skripsi ini.
5. Kepada seseorang yang pernah menjadi bagian yang sangat berarti dalam perjalanan hidup penulis. Terima kasih atas pendewasaan yang diajarkan kepada penulis, segala dukungan, perhatian, semangat, dan kebaikan yang pernah diberikan. Tanpa disadari, kehadiran, kepergian dan kata-kata yang pernah diberikan menjadi salah satu motivasi bagi penulis untuk terus berusaha, sabar, bertahan, menerima, dan menjadi pribadi yang lebih baik. Mungkin tidak semua perjalanan harus berakhir dengan tetap berjalan bersama, tetapi setiap hal yang pernah diberikan akan selalu di ingat dan menjadi pengalaman terbaik bagi penulis. Terima kasih telah pernah menjadi bagian dari perjalanan yang turut membawa penulis sampai pada titik ini. Pada akhirnya setiap orang ada masanya dan setiap masa ada orangnya.
6. Kepada diri sendiri, terima kasih telah bertahan, berusaha, dan tidak menyerah meskipun perjalanan ini tidak selalu mudah. Terima kasih telah tetap melangkah ketika merasa lelah, tetap mencoba ketika merasa ragu, dan mampu menyelesaikan apa yang telah dimulai. Semoga pencapaian ini menjadi pengingat bahwa setiap perjuangan, sekecil apa pun, akan memiliki arti ketika dijalani dengan kesabaran dan ketekunan.
7. Kepada semua pihak yang tidak dapat penulis sebutkan satu per satu. Terima kasih atas setiap doa, bantuan, dukungan, dan kebaikan yang telah diberikan selama proses perkuliahan hingga penyelesaian skripsi ini. Semoga segala kebaikan yang diberikan mendapatkan balasan terbaik.

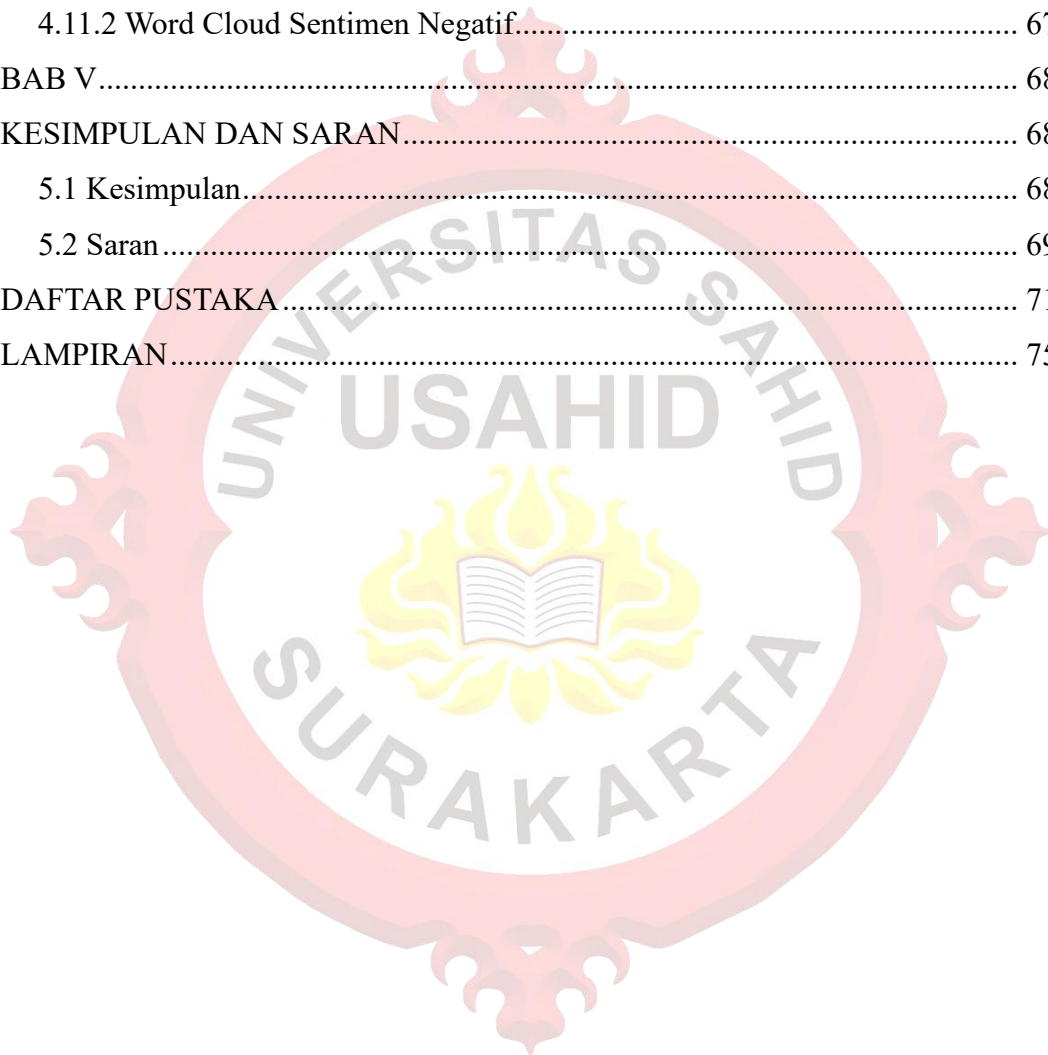
DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
SURAT PERNYATAAN ORISINALITAS KARYA ILMIAH	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	v
KATA PENGANTAR.....	vi
LEMBAR PERSEMBAHAN	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
ABSTRAK	xv
ABSTRACT.....	xvi
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Perumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan dan Manfaat Penelitian	6
1.4.1 Tujuan.....	6
1.4.2 Manfaat.....	6
BAB II.....	8
TINJAUAN PUSTAKA	8
2.1 Penelitian Terdahulu.....	8
2.2 Landasan Teori	12
2.2.1 Financial Technology	12
2.2.2 Kredivo	12
2.2.3 Google Play Store.....	12
2.2.4 Ulasan.....	12

2.2.5 Analisis Sentimen	13
2.2.6 Text Mining	13
2.2.7 Natural Language Processing	13
2.2.8 Preprocessing.....	13
2.2.9 Pelabelan Data	14
2.2.10 TF-IDF (Term Frequency-Inverse Document Frequency)	15
2.2.11 Splitting Data.....	15
2.2.12 K-Fold Cross Validation	15
2.2.13 SMOTE (Synthetic Minority Oversampling Technique)	16
2.2.14 Naive Bayes.....	16
2.2.15 Logistic Regression	17
2.2.16 Cofusion Matrix	17
2.2.17 Indo-Normalizer	19
2.2.18 Google-Play-Scraper	19
2.2.19 Sastrawi	19
2.2.20 Google Colaboratory	19
2.3 Kerangka Pemikiran.....	20
2.3.1 Analisis Masalah.....	21
2.3.2 Studi Literatur.....	21
2.3.3 Pengumpulan Data	21
2.3.4 Preprocessing Data	22
2.3.5 Pelabelan Berbasis Rating.....	22
2.3.6 Ekstraksi Fitur TF-IDF	23
2.3.7 Pembagian Data (Train-Test Split)	23
2.3.8 K-Fold Cross Validation	23
2.3.9 Synthetic Minority Oversampling Technique (SMOTE)	23
2.3.10 Pelatihan Model.....	24
2.3.11 Pengujian Model.....	24
2.3.12 Evaluasi	24
2.3.13 Hasil Penelitian.....	25
BAB III	26

METODOLOGI PENELITIAN.....	26
3.1 Metode Penelitian.....	26
3.2 Sumber Data.....	26
3.3 Alat dan Bahan Penelitian	27
3.3.1 Perangkat Keras (Hardware)	27
3.3.2 Perangkat Lunak (Software)	28
3.4 Alur Penelitian.....	29
3.4.1 Pengumpulan Data	30
3.4.2 Preprocessing.....	31
3.4.3 Pelabelan Data (Data Labeling)	32
3.4.4 Ekstraksi Fitur TF-IDF.....	33
3.4.5 Splitting Data.....	34
3.4.6 K-Fold Cross Validation	35
3.4.7 Synthetic Minority Oversampling Technique (SMOTE)	36
3.4.8 Penerapan Algoritma	36
3.4.9 Pengujian Data Testing	37
3.4.10 Evaluasi Model.....	38
BAB IV	40
HASIL DAN PEMBAHASAN	40
4.1 Deskripsi Penelitian	40
4.2 Hasil Pengumpulan Data.....	40
4.3 Hasil Preprocessing.....	42
4.4 Hasil Pelabelan Berbasis Rating	43
4.5 Hasil Ekstraksi Fitur TF-IDF	45
4.6 Hasil Pembagian Data (Splitting Data 80:20)	47
4.7 Hasil Implementasi K-Fold Cross Validation	47
4.8 Hasil Penanganan Ketidakseimbangan Data (SMOTE)	48
4.9 Hasil Pelatihan Model.....	50
4.9.1 Naive Bayes.....	50
4.9.2 Logistic Regression	53
4.9.3 Perbandingan Model.....	56

4.10 Hasil Pengujian Model (Data Testing 20%).....	58
4.10.1 Hasil Pengujian Naive Bayes	58
4.10.2 Hasil Pengujian Logistic Regression.....	61
4.10.3 Perbandingan Pengujian Model.....	63
4.11 Visualisasi Word Cloud	65
4.11.1 Word Cloud Setimen Positif.....	66
4.11.2 Word Cloud Sentimen Negatif.....	67
BAB V.....	68
KESIMPULAN DAN SARAN.....	68
5.1 Kesimpulan.....	68
5.2 Saran.....	69
DAFTAR PUSTAKA.....	71
LAMPIRAN.....	75



DAFTAR GAMBAR

Gambar 2. 1 Confusion Matrix	17
Gambar 2. 2 Kerangka Pemikiran	20
Gambar 3. 1 Diagram Alur Penelitian.....	29
Gambar 4. 1 Tabel Perubahan Pada Setiap Preprocessing.....	43
Gambar 4. 2 Distribusi Label Sentimen.....	45
Gambar 4. 3 Confusion Matrix Naive Bayes	52
Gambar 4. 4 Confusion Matrix Logistic Regression	55
Gambar 4. 5 Perbandingan Akurasi Model.....	57
Gambar 4. 6 Confusion Matrix Naive Bayes Data Testing 20%.....	60
Gambar 4. 7 Confusion Matrix Logistic Regression Data Testing 20%.....	62
Gambar 4. 8 Perbandingan Performa Model Pada Data Testing 20%.....	64
Gambar 4. 9 Word Cloud Sentimen Positif	66
Gambar 4. 10 Word Cloud Sentimen Negatif.....	67

DAFTAR TABEL

Tabel 3. 1 Sumber Data.....	27
Tabel 3. 2 Komponen Perangkat Keras.....	28
Tabel 3. 3 Komponen Perangkat Lunak.....	28
Tabel 3. 4 Contoh Ulasan Kredivo di Google Play Store	30
Tabel 3. 5 Aturan Perlabelan Data	33
Tabel 3. 6 Visualisasi K-Fold Cross Validation.....	35
Tabel 4. 1 Pengumpulan Data Ulasan Aplikasi Kredivo	41
Tabel 4. 2 Hasil Perlabelan Verifikasi Manual Rating 3	44
Tabel 4. 3 Ringkasan Parameter dan Output Ekstraksi Fitur TF-IDF.....	46
Tabel 4. 4 Bobot TF-IDF pada Dokumen Pertama	46
Tabel 4. 5 Rincian Pembagian Dataset Penelitian	47
Tabel 4. 6 Hasil Implementasi K-Fold Cross Validation.....	48
Tabel 4. 7 Perubahan Jumlah Sebelum Dan Sesudah SMOTE.....	49
Tabel 4. 8 Performa Evaluasi Model Naive Bayes pada Setiap Fold	51
Tabel 4. 9 Rata-Rata Kinerja Naive Bayes Berdasarkan Kelas Sentimen	51
Tabel 4. 10 Hasil Evaluasi Model pada Setiap Fold	54
Tabel 4. 11 Evaluasi Kinerja Model Berdasarkan Kelas Sentimen	54
Tabel 4. 12 Perbandingan Nilai Rata-Rata Evaluasi Model	56
Tabel 4. 13 Performa Akhir Model Naive Bayes pada Data Testing.....	59
Tabel 4. 14 Kinerja Naive Bayes Per Kelas Data Testing	59
Tabel 4. 15 Performa Akhir Model Logistic Regression Data Testing.....	61
Tabel 4. 16 Kinerja Logistic Regression Per Kelas Data Testing.....	62
Tabel 4. 17 Perbandingan Performa Akhir Model Data Testing (20%)	64

ABSTRAK

Perkembangan pesat industri Financial Technology (Fintech) Buy Now Pay Later (BNPL) di Indonesia mendorong tingginya akumulasi ulasan pengguna pada aplikasi Kredivo di Google Play Store. Ulasan tersebut memuat data opini yang sangat berharga terkait tingkat kepuasan, keandalan sistem, maupun keluhan layanan, namun volumenya yang besar menyulitkan evaluasi secara manual. Penelitian ini bertujuan untuk mengidentifikasi sentimen publik serta membandingkan performa algoritma Multinomial Naive Bayes dan Logistic Regression dalam mengklasifikasikan sentimen ulasan pengguna Kredivo. Dataset sebanyak 17.212 ulasan terfilter periode Januari 2021 hingga Januari 2026 diproses melalui tahapan preprocessing (cleansing, case folding, tokenizing, normalisasi kata Indo-Normalizer, stopword removal, dan stemming Sastrawi) hingga menghasilkan 10.798 data ulasan bersih. Pelabelan otomatis berbasis rating dengan verifikasi manual pada rating 3 menghasilkan distribusi sentimen yang tidak seimbang, yaitu 80,55% positif dan 19,45% negatif. Pembobotan kata menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) menghasilkan 4.558 fitur kata. Untuk mengatasi ketidakseimbangan kelas tanpa memicu data leakage, teknik Synthetic Minority Oversampling Technique (SMOTE) diterapkan secara terisolasi pada data latih di setiap fold menggunakan 5-Fold Cross Validation. Hasil eksperimen menunjukkan Logistic Regression unggul secara konsisten dibanding Naive Bayes dengan menggapai rata-rata akurasi sebesar 90,14%, precision 91,43%, recall 90,14%, dan F1-score 90,52% pada 5-Fold Cross Validation. Pada pengujian akhir menggunakan 20% data uji murni (2.160 ulasan), Logistic Regression kembali membuktikan daya generalisasi terbaik tanpa overfitting dengan akurasi 90,42% dan F1-score 90,74%. Analisis Word Cloud mengungkapkan bahwa kepuasan pengguna didominasi oleh efisiensi pencairan dana dan kemudahan aplikasi, sedangkan keluhan terfokus pada penolakan limit transaksi mendadak serta mekanisme penagihan.

Kata Kunci: Analisis Sentimen, Kredivo, Naive Bayes, Logistic Regression, Fintech.

ABSTRACT

The rapid growth of the Buy Now Pay Later (BNPL) Financial Technology (Fintech) industry in Indonesia has led to a massive accumulation of user reviews for the Kredivo application on the Google Play Store. These reviews contain valuable opinion data regarding satisfaction levels, system reliability, and service complaints; however, the sheer volume makes manual evaluation difficult. This study aims to identify public sentiment and compare the performance of Multinomial Naive Bayes and Logistic Regression algorithms in classifying Kredivo user review sentiments. A dataset of 17,212 reviews filtered for the period from January 2021 to January 2026 underwent preprocessing stages (cleansing, case folding, tokenizing, Indo-Normalizer word normalization, stopword removal, and Sastrawi stemming), resulting in 10,798 clean review entries. Automated rating-based labeling, with manual verification for 3-star ratings, yielded an imbalanced sentiment distribution: 80.55% positive and 19.45% negative. Word weighting using Term Frequency-Inverse Document Frequency (TF-IDF) generated 4,558 word features. To address class imbalance without causing data leakage, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training data within each fold during 5-Fold Cross-Validation. Experimental results demonstrate that Logistic Regression consistently outperformed Naive Bayes, achieving an average accuracy of 90.14%, precision of 91.43%, recall of 90.14%, and an F1-score of 90.52% across the 5-Fold Cross-Validation. In the final test on a hold-out set of 20% of the data (2,160 reviews), Logistic Regression again showed strong generalization without overfitting, achieving 90.42% accuracy and a 90.74% F1-score. Word Cloud analysis shows that user satisfaction is driven mainly by efficient fund disbursement and ease of use, while complaints focus on sudden transaction limit rejections and collection mechanisms.

Keywords: Sentiment Analysis, Kredivo, Naive Bayes, Logistic Regression, Fintech.

