



**ANALISIS SENTIMEN PENGGUNA APLIKASI *OCTO
MOBILE BY CIMB NIAGA* DI *GOOGLE PLAY STORE*
MENGUNAKAN *LATENT DIRICHLET ALLOCATION (LDA)*
DAN *SUPPORT VECTOR MACHINE (SVM)***

LAPORAN TUGAS AKHIR

**SATYA KRISTIAN PERWIRA SIMAMORA
41821010117**

**TUTI FARADILLA MASTATO
41821010088**

**MUHAMMAD RAFLY AZIZ
41821010054**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**ANALISIS SENTIMEN PENGGUNA APLIKASI *OCTO
MOBILE BY CIMB NIAGA* DI *GOOGLE PLAY STORE*
MENGUNAKAN *LATENT DIRICHLET ALLOCATION (LDA)*
DAN *SUPPORT VECTOR MACHINE (SVM)***

LAPORAN TUGAS AKHIR

**SATYA KRISTIAN PERWIRA SIMAMORA
41821010117**

**TUTI FARADILLA MASTATO
41821010088**

**MUHAMMAD RAFLY AZIZ
41821010054**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2025

HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Satya Kristian Perwira Simamora
NIM : 41821010117
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Analisis Sentimen Pengguna Aplikasi Octo Mobile by CIMB Niaga di Google Play Store Menggunakan Latent Dirichlet Allocation (LDA) dan Support Vector Machine (SVM)

Menyatakan bahwa Laporan Aplikatif/Tugas Akhir/Jurnal/Media Ilmiah ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 9 Juli 2025

UNIVERSITAS
MERCU BUANA



Satya Kristian Perwira Simamora

LEMBAR PENGESAHAN

Nama Mahasiswa (1) : Satya Kristian Perwira Simamora
NIM : 41821010117
Nama Mahasiswa (2) : Tuti Faradilla Mastato
NIM : 41821010088
Nama Mahasiswa (3) : Muhammad Rafly Aziz
NIM : 41821010045
Judul Tugas Akhir : Analisis Sentimen Pengguna Aplikasi Octo Mobile
by CIMB Niaga di Google Play Store Menggunakan
Latent Dirichlet Allocation (LDA) dan Support
Vector Machine (SVM)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

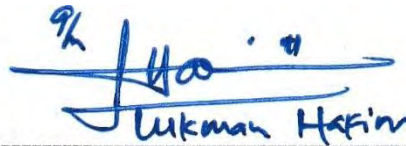
Jakarta, 9 Juli 2025

Menyetujui

Pembimbing : Yuwan Jumaryadi, S.Kom, M.Kom ()
M.Kom
NIDN : 0319078704
Ketua Penguji : Nur Ani, ST, MMSI ()
NIDN : 0310117801
Penguji 1 : Misni, S.Kom, M.Kom ()
NIDN : 0413046802
Penguji 2 : Fajar Masya, Ir. MMSI ()
NIDN : 0313036701

Mengetahui,


Dr. Bambang Jokonowo, S.Si., M.T.I.
Dekan Fakultas Ilmu Komputer


Ka.Prodi Sistem Informasi

KATA PENGANTAR

Segala Puji dan Syukur kita panjatkan kehadirat Allah yang telah memberikan Rahmat dan Karunianya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “Analisis Sentimen Pengguna Aplikasi *Octo Mobile by CIMB Niaga* di *Google Play Store* Menggunakan *Latent Dirichlet Allocation (LDA)* dan *Support Vector Machine (SVM)*” sebagai salah satu syarat Program Sarjana (S1) pada Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Mercu Buana.

Penulis menyadari bahwa penulisan ini tidak dapat terselesaikan tanpa bantuan dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis ingin mengucapkan penghargaan kepada:

1. Kedua orang tua, yang senantiasa memberikan doa, dukungan, serta motivasi sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan baik.
2. Dosen pembimbing, Bapak Yuwan Jumaryadi, S.Kom, MM, M.Kom, yang telah memberikan arahan, dukungan, dan motivasi dengan tulus untuk membantu penulis menyelesaikan Tugas Akhir ini.
3. Ibu Dr. Ruci Meiyanti, M.Kom, selaku Ketua Program Studi Sistem Informasi Universitas Mercu Buana.
4. Seluruh dosen dan staf Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Mercu Buana yang telah membantu penulis dalam proses penyusunan Tugas Akhir ini.

Penulis memiliki harapan besar agar Tugas Akhir ini dapat memberikan kontribusi terhadap pengembangan ilmu pengetahuan, khususnya di bidang Sistem Informasi. Semoga karya ini dapat bermanfaat bagi para pembaca sebagai referensi untuk penelitian yang lebih baik di masa mendatang. Penulis juga dengan senang hati menerima kritik dan saran demi perbaikan di masa yang akan datang.

Jakarta, 9 Juli 2025

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buanas saya yang bertanda tangan di bawah ini:

Nama : Satya Kristian Perwira Simamora
NIM : 41821010117
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Analisis Sentimen Pengguna Aplikasi Octo Mobile by CIMB Niaga
di Google Play Store Menggunakan Latent Dirichlet Allocation
(LDA) dan Support Vector Machine (SVM)

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana Hak Bebas Royalti Non-Eksklusif (Nonexclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 9 Juli 2025

Yang menyatakan,



Satya Kristian Perwira Simamora

ABSTRAK

Nama Mahasiswa : Satya Kristian Perwira Simamora
NIM : 41821010117
Program Studi : Sistem Informasi
Judul Tugas Akhir : Analisis Sentimen Pengguna Aplikasi *Octo Mobile by CIMB Niaga* di *Google Play Store* Menggunakan *Latent Dirichlet Allocation* (LDA) dan *Support Vector Machine* (SVM)
Pembimbing : Yuwan Jumaryadi, S.Kom, MM, M.Kom

Penurunan rating aplikasi *Octo Mobile by CIMB Niaga* di *Google Play Store* menunjukkan adanya ketidakpuasan pengguna, namun permasalahan utamanya adalah belum teridentifikasinya penyebab spesifik di balik penurunan tersebut. Penelitian ini menawarkan solusi dengan menerapkan pendekatan analisis gabungan: menggunakan *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen secara akurat, dan *Latent Dirichlet Allocation* (LDA) untuk mengekstrak topik utama sebagai akar penyebab keluhan. Metode yang dijalankan meliputi perbandingan enam algoritma pada 22.389 ulasan melalui validasi silang 10-fold, yang dilanjutkan dengan pemodelan topik. Hasilnya, SVM terbukti sebagai model terbaik dengan akurasi 95%, dan LDA berhasil mendiagnosis keluhan utama pada topik “Aksesibilitas Akun” dan “Fungsionalitas Aplikasi” yang frekuensinya melonjak pada versi 3.0 dan 3.1. Kombinasi metode ini terbukti efektif memberikan diagnosis masalah yang jelas dan wawasan yang dapat ditindaklanjuti oleh pengembang.

Kata Kunci: Analisis Sentimen, *Support Vector Machine*, *Latent Dirichlet Allocation*, Ulasan Pengguna, *Octo Mobile by CIMB Niaga*

ABSTRACT

Name : Satya Kristian Perwira Simamora
NIM : 41821010117
Study Program : Sistem Informasi
Thesis Title : Analisis Sentimen Pengguna Aplikasi *Octo Mobile by CIMB Niaga* di *Google Play Store* Menggunakan *Latent Dirichlet Allocation (LDA)* dan *Support Vector Machine (SVM)*
Counsellor : Yuwan Jumaryadi, S.Kom, MM, M.Kom

The decline in the rating of the Octo Mobile by CIMB Niaga application on the Google Play Store indicates user dissatisfaction, but the main problem is that the specific cause behind the decline has not been identified. This research offers a solution by applying a combined analysis approach: using Support Vector Machine (SVM) to accurately classify sentiment, and Latent Dirichlet Allocation (LDA) to extract key topics as the root cause of complaints. The method involved comparing the six algorithms on 22,389 reviews through 10-fold cross-validation, followed by topic modeling. As a result, SVM proved to be the best model with 95% accuracy, and LDA successfully diagnosed major complaints on the topics “Account Accessibility” and “App Functionality” whose frequency spiked in versions 3.0 and 3.1. This combination of methods proved effective in providing clear problem diagnosis and actionable insights for developers.

Keywords: Sentiment Analysis, Support Vector Machine, Latent Dirichlet Allocation, User Reviews, Octo Mobile by CIMB Niaga

DAFTAR ISI

HALAMAN JUDUL	i
DAN <i>SUPPORT VECTOR MACHINE</i> (SVM)	i
HALAMAN PENYATAAN KARYA SENDIRI	ii
LEMBAR PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	iv
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Permasalahan.....	6
1.3 Tujuan Penelitian.....	6
1.4 Batasan Masalah.....	6
1.5 Manfaat Penelitian	7
1.6. Sistematika Penulisan.....	8
BAB II TINJAUAN PUSTAKA	9
2.1 Teori/ Konsep Terkait.....	9
2.1.1 Analisis Sentimen.....	9
2.1.2 Web scraping	10

2.1.3	<i>Preprocessing Data dalam Analisis Sentimen</i>	11
2.1.4	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	12
2.1.5	<i>Validasi Silang 10-Fold (10-Fold Cross Validation)</i>	14
2.1.6	<i>Evaluasi Hasil</i>	15
2.1.7	<i>Support Vector Machine (SVM)</i>	18
2.1.8	<i>Latent Dirichlet Allocation (LDA)</i>	19
2.2	<i>Penelitian Terdahulu</i>	20
2.3	<i>Analisis Literature Review</i>	37
BAB III METODE PENELITIAN		39
3.1	<i>Deskripsi Sumber Data</i>	39
3.2	<i>Teknik Pengumpulan Data</i>	39
3.3	<i>Diagram Alir Penelitian</i>	40
3.4	<i>Jadwal Penelitian</i>	43
BAB IV HASIL DAN PEMBAHASAN		44
4.1	<i>Pengumpulan Data</i>	44
4.2	<i>Preprocessing Data</i>	46
4.2.1	<i>Data Cleansing</i>	48
4.2.2	<i>Case Folding</i>	49
4.2.3	<i>Normalization</i>	50
4.2.4	<i>Stopwords Removal</i>	50
4.2.5	<i>Stemming</i>	51
4.2.6	<i>Tokenization</i>	51
4.3	<i>Pelabelan Manual (Manual Labelling)</i>	53
4.4	<i>Modelling</i>	54
4.4.1	<i>Pembobotan TF-IDF</i>	55
4.4.2	<i>10-Fold Cross Validation</i>	57

4.4.3 <i>Training-Testing</i>	58
4.4.4 Evaluasi Model	59
4.5 Hasil Evaluasi dan Perbandingan Model	63
4.6 Penerapan Algoritma	77
4.7 Implementasi <i>Latent Dirichlet Allocation</i> (LDA)	79
BAB V PENUTUP	85
5.1 Kesimpulan	85
5.2 Saran	86
DAFTAR PUSTAKA	87
LAMPIRAN	94
Lampiran 1.1 Lembar Konsultasi Bimbingan Satya Kristian Perwira	94
Lampiran 1.2 Lembar Konsultasi Bimbingan Tuti Faradilla Mastato	95
Lampiran 1.3 Lembar Konsultasi Bimbingan Muhammad Rafly Aziz	96
Lampiran 2.1 Biodata Satya Kristian Perwira	97
Lampiran 2.2 Biodata Tuti Faradilla Mastato	98
Lampiran 2.3 Biodata Muhammad Rafly Aziz	99
Lampiran 2.4 Hasil Cek Uji Plagiasi (Turnitin)	100

DAFTAR TABEL

Tabel 2. 1 <i>Confusion matrix</i>	15
Tabel 4. 1 Contoh Hasil Pelabelan Manual	54
Tabel 4. 2 Jumlah Data Skenario Penelitian.....	55
Tabel 4. 3 Evaluasi Model <i>Naïve Bayes</i> (%).....	65
Tabel 4. 4 Evaluasi Model KNN (%).....	67
Tabel 4. 5 Evaluasi Model <i>Decision Tree</i> (%).....	69
Tabel 4. 6 Evaluasi Mode <i>Logistic Regression</i> (%)	70
Tabel 4. 7 Evaluasi Model SVM (%).....	72
Tabel 4. 8 Evaluasi Model <i>Random Forest</i> (%).....	74



DAFTAR GAMBAR

Gambar 1. 1 Rata-Rata Tahunan <i>Rating</i> Aplikasi <i>Octo Mobile</i> by CIMB Niaga pada <i>Google Play Store</i> (2019-2024).....	2
Gambar 1. 2 Rata-Rata <i>Rating</i> per Versi Mayor Aplikasi <i>Octo Mobile</i> by CIMB Niaga pada <i>Google Play Store</i>	3
Gambar 3. 1 Sampel Data	40
Gambar 3. 2 Diagram Alir Penelitian.....	40
Gambar 3. 3 Jadwal Penelitian.....	43
Gambar 4. 1 Instalasi <i>Library</i> untuk <i>Scraping</i> Data	44
Gambar 4. 2 Proses <i>Scraping</i> dan Penyimpanan Ulasan <i>Google Play</i>	45
Gambar 4. 3 Hasil <i>Scraping</i>	46
Gambar 4. 4 Instalasi <i>Library</i> Sastrawi.....	46
Gambar 4. 5 Inisialisasi <i>Library</i> dan Pembacaan Dataset.....	47
Gambar 4. 6 Load Resource dan Pembersihan Awal Data.....	48
Gambar 4. 7 Data <i>Cleansing</i>	49
Gambar 4. 8 <i>Case Folding</i>	49
Gambar 4. 9 <i>Normalization</i>	50
Gambar 4. 10 <i>Stopwords Removal</i>	51
Gambar 4. 11 <i>Stemming</i>	51
Gambar 4. 12 Final <i>Clean</i> dan Tokenisasi	52
Gambar 4. 13 Hasil <i>Preprocessing</i> Data Ulasan	52
Gambar 4. 14 Jumlah Kata per Tahapan <i>Preprocessing</i>	53
Gambar 4. 15 Impor <i>Library</i> dan Pemuatan Dataset	55
Gambar 4. 16 Pembobotan TF-IDF	56
Gambar 4. 17 <i>10-Fold Cross validation</i>	57
Gambar 4. 18 <i>Training-Testing</i>	58
Gambar 4. 19 Perhitungan Metrik.....	59
Gambar 4. 20 <i>Metric 10-fold Visualization</i>	60
Gambar 4. 21 <i>Report & Confusion Matrix</i>	60
Gambar 4. 22 Komparasi Akurasi.....	61
Gambar 4. 23 <i>ROC Curves</i>	62

Gambar 4. 24 Hasil <i>Metric 10-fold Visualization</i>	63
Gambar 4. 25 <i>Confusion Matrix: Naïve Bayes</i>	65
Gambar 4. 26 <i>Confusion Matrix: KNN</i>	66
Gambar 4. 27 <i>Confusion Matrix: Decision Tree</i>	68
Gambar 4. 28 <i>Confusion Matrix: Logistic Regression</i>	70
Gambar 4. 29 <i>Confusion Matrix: SVM</i>	71
Gambar 4. 30 <i>Confusion Matrix: Random Forest</i>	73
Gambar 4. 31 Hasil Komparasi Akurasi	75
Gambar 4. 32 ROC-AUC Curves.....	76
Gambar 4. 33 Penerapan Algoritma (SVM).....	77
Gambar 4. 34 Visualisasi Penerapan Algoritma (SVM)	78
Gambar 4. 35 Distribusi Label Berdasarkan Sumber.....	79
Gambar 4. 36 Implementasi <i>Latent Dirichlet Allocation</i> (LDA).....	80
Gambar 4. 37 Visualisasi Topik Utama Ulasan Negatif dan Positif	81
Gambar 4. 38 Sebaran Topik per Versi Aplikasi	83
Gambar 4. 39 Visualisasi Frekuensi Topik per <i>App Version</i>	84

DAFTAR LAMPIRAN

Lampiran 1.1 Lembar Konsultasi Bimbingan Satya Kristian Perwira.....	94
Lampiran 1.2 Lembar Konsultasi Bimbingan Tuti Faradilla Mastato	95
Lampiran 1.3 Lembar Konsultasi Bimbingan Muhammad Rafly Aziz	96
Lampiran 2.1 Biodata Satya Kristian Perwira.....	97
Lampiran 2.2 Biodata Tuti Faradilla Mastato	98
Lampiran 2.3 Biodata Muhammad Rafly Aziz	99
Lampiran 2.4 Hasil Cek Uji Plagiasi (Turnitin).....	100

