

**ANALISIS SENTIMEN ULASAN APLIKASI HALO BCA
DENGAN METODE LSTM, NAIVE BAYES DAN SVM**



SKRIPSI

JONATHAN JASON JUSMAN

00000054182

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**ANALISIS SENTIMEN ULASAN APLIKASI HALO BCA
DENGAN METODE LSTM, NAIVE BAYES DAN SVM**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

JONATHAN JASON JUSMAN

00000054182

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Jonathan Jason Jusman

Nomor Induk Mahasiswa : 00000054182

Program Studi : Sistem Informasi

Proposal Skripsi dengan judul:

ANALISIS SENTIMEN ULASAN APLIKASI HALO BCA DENGAN METODE LSTM, NAIVE BAYES DAN SVM

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 19 Juni 2025



Jonathan Jason Jusman

HALAMAN PENGESAHAN

Skripsi dengan judul
ANALISIS SENTIMEN ULASAN APLIKASI HALO BCA DENGAN
METODE LSTM, NAÏVE BAYES DAN SVM

Oleh

Nama : Jonathan Jason Jusman
NIM : 00000054182
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

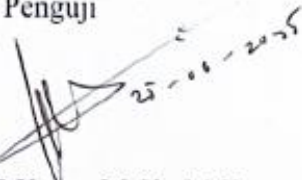
Telah diujikan pada hari Senin, 16 Juni 2025
Pukul 15.00 s.d 17.00 dan dinyatakan
LULUS
Dengan susunan penguji sebagai berikut.

Ketua Sidang

Penguji

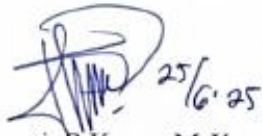

Dr. Santo Fernandi Wijaya, S.Kom., M.M.
081435

Pembimbing


Rudi Sutomo, S.Kom., M.Si., M.Kom.
0222057501


Johan Setiawan, S.Kom., M.M., MBA
327106402

Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.Kom., M.Kom.
0313058001

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Jonathan Jason Jusman
NIM : 00000054182
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Ulasan Aplikasi
Halo BCA dengan Metode LSTM,
Naïve Bayes dan SVM.

Menyatakan dengan sesungguhnya bahwa saya bersedia* (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara R Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 19 Juni 2025



Jonathan Jason Jusman

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji syukur dipanjatkan ke hadirat Tuhan Yang Maha Esa berkat-Nya sehingga dapat menyelesaikan skripsi yang berjudul "Analisis Sentimen Ulasan Aplikasi Halo BCA dengan Metode LSTM, Naïve Bayes, dan SVM". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dan bertujuan untuk menganalisis persepsi pengguna terhadap aplikasi Halo BCA melalui pendekatan *machine learning*. Semoga hasil skripsi ini bermanfaat bagi pengembangan keilmuan dan praktik analisis sentimen ke depannya dan mengucapkan terima kasih kepada:

1. Bapak Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Bapak Johan Setiawan, S.Kom., MM., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini dapat menjadi sumber informasi dan inspirasi bagi pembaca.

Tangerang, 19 Juni 2025



Jonathan Jason Jusman

ANALISIS SENTIMEN ULASAN APLIKASI HALO BCA DENGAN METODE LSTM, NAÏVE BAYES DAN SVM

Jonathan Jason Jusman

ABSTRAK

Aplikasi Halo BCA merupakan salah satu layanan perbankan digital yang banyak digunakan oleh masyarakat Indonesia untuk berbagai kebutuhan transaksi dan layanan pelanggan. Seiring meningkatnya penggunaan aplikasi ini, ulasan pengguna yang terdapat di Google Play Store menjadi sumber informasi penting dalam mengevaluasi kualitas layanan. Analisis sentimen terhadap ulasan tersebut dapat memberikan wawasan berharga mengenai persepsi pengguna yang dapat dimanfaatkan oleh tim pengembang aplikasi dalam meningkatkan mutu dan responsivitas layanan digital perbankan.

Penelitian ini menggunakan tiga metode klasifikasi sentimen, yaitu Long Short-Term Memory (LSTM), Naïve Bayes, dan Support Vector Machine (SVM). Proses dimulai dengan pengumpulan data ulasan dari Google Play Store, dilanjutkan dengan tahapan *pre-processing* seperti *case folding*, tokenisasi, *stopword removal*, dan *stemming*. Representasi fitur dilakukan menggunakan TF-IDF untuk model SVM dan Naïve Bayes, serta tokenisasi dan padding untuk LSTM. Untuk mengatasi ketidakseimbangan data antara sentimen positif dan negatif, diterapkan teknik SMOTE (Synthetic Minority Over-sampling Technique) sebelum proses pelatihan model. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score.

Hasil penelitian menunjukkan bahwa penerapan SMOTE memberikan peningkatan performa pada seluruh model, khususnya pada LSTM yang sebelumnya kurang optimal. Model SVM mencatat akurasi tertinggi sebesar 82%, diikuti oleh LSTM dengan 81%, dan Naïve Bayes sebesar 77%. SVM dinilai sebagai model paling konsisten dan seimbang dalam mengklasifikasikan kedua kelas sentimen. Berdasarkan hasil ini, model SVM direkomendasikan untuk diimplementasikan dalam sistem pemantauan ulasan secara otomatis. Penelitian ini diharapkan dapat menjadi dasar pengambilan keputusan strategis dalam pengembangan layanan aplikasi Halo BCA yang lebih adaptif terhadap opini dan kebutuhan pengguna.

Kata kunci: Aplikasi Halo BCA, Analisis Sentimen, LSTM, Naïve Bayes, SVM, SMOTE, Ulasan Pengguna

Sentiment Analysis of Halo BCA Application Reviews

Using LSTM, Naïve Bayes and SVM

Jonathan Jason Jusman

ABSTRACT (English)

Halo BCA is one of the most widely used digital banking applications in Indonesia, serving various user needs such as financial transactions and customer services. With the increasing usage of the application, user reviews on the Google Play Store have become a valuable source of insights for evaluating service quality. Sentiment analysis of these reviews can provide meaningful information about user perceptions, which can be utilized by application developers to improve the responsiveness and quality of digital banking services.

This study employs three sentiment classification methods: Long Short-Term Memory (LSTM), Naïve Bayes, and Support Vector Machine (SVM). The research process begins with collecting user reviews from the Google Play Store, followed by a series of preprocessing steps, including case folding, tokenization, stopword removal, and stemming. Feature representation is carried out using TF-IDF for SVM and Naïve Bayes models, and tokenization with padding for the LSTM model. To address the class imbalance between positive and negative sentiments, the SMOTE (Synthetic Minority Over-sampling Technique) method is applied prior to model training. The models are evaluated using metrics such as accuracy, precision, recall, and F1-score.

The results indicate that SMOTE significantly improves the performance of all models, particularly LSTM, which previously underperformed. The SVM model achieved the highest accuracy at 82%, followed by LSTM at 81%, and Naïve Bayes at 77%. SVM is considered the most consistent and balanced model in classifying both sentiment classes. Based on these findings, SVM is recommended for implementation in an automated review monitoring system. This research is expected to serve as a valuable reference for strategic decision-making in enhancing the Halo BCA application's ability to adapt to user feedback and improve service quality.

Keywords: Halo BCA Application, Sentiment Analysis, LSTM, Naïve Bayes, SVM, SMOTE, User Reviews

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
KATA PENGANTAR	v
ABSTRAK	vi
<i>ABSTRACT (English)</i>	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	5
1.4.1 Tujuan Penelitian	5
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Penelitian Terdahulu	7
2.2 Teori Penelitian	16
2.2.1 Analisis Sentimen	16
2.2.2 Aplikasi <i>Mobile Banking</i>	16
2.2.3 Ulasan Pengguna	16
2.2.4 Metode Support Vector Machine	16
2.2.5 Metode Naïve Bayes	18
2.2.6 Metode LSTM	19
2.2.7 Teknik Preprocessing Data	21
2.2.8 Synthetic Minority Oversampling Technique (SMOTE)	22

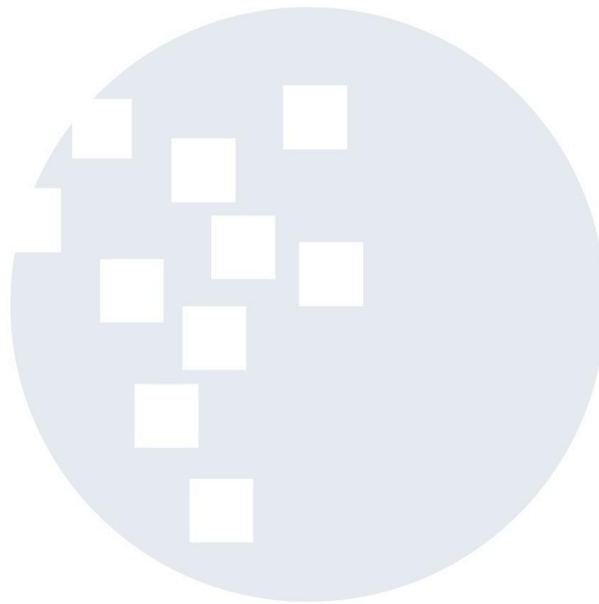
2.3	Framework dan Algoritma Penelitian	22
2.3.1	CRISP-DM (Cross-Industry Standard Process for Data Mining) 22	
2.3.2	Support Vector Machine	23
2.3.3	Naïve Bayes	23
2.3.4	Long Short – Term Memory (LSTM)	24
2.4	Tools dan Software Penelitian	26
2.4.1	Python	26
2.4.2	Jupyter Notebook.....	26
2.4.3	Scikit-Learn	26
2.4.4	NLTK (Natural Language Toolkit)	26
BAB III	METODOLOGI PENELITIAN	27
3.1	Gambaran Umum Objek Penelitian.....	27
3.2	Metode Penelitian.....	27
3.2.1	Alur Penelitian.....	27
3.2.2	Metode Data Mining	29
3.3	Teknik Pengumpulan Data	31
3.3.1	Populasi dan Sampel.....	31
3.3.2	Periode Pengambilan Data	32
3.4	Variabel Penelitian.....	32
3.4.1	Variabel Independen (X).....	32
3.4.2	Variabel Dependen (Y)	32
3.5	Teknik Analisis Data.....	33
BAB IV	ANALISIS DAN HASIL PENELITIAN	36
4.1	Business Understanding	36
4.2	Data Understanding	37
4.3	Data Preparation.....	40
4.3.1	Drop <i>Duplicate</i> Data.....	41
4.3.2	<i>Cleaning</i> Data	42
4.3.3	<i>Preprocessing</i> Data	43
4.3.4	Labelling	46
4.3.5	TF-IDF (Term Frequency–Inverse Document Frequency).....	49

4.3.6	Tokenisasi dan Padding.....	53
4.4	Modelling	56
4.4.1	Word Cloud.....	56
4.4.2	SVM	57
4.4.3	Naïve Bayes.....	60
4.4.4	LSTM	62
4.5	Evaluation	65
4.5.1	SVM	65
4.5.2	Naive Bayes.....	69
4.5.3	LSTM	73
4.6	Deployment.....	76
4.7	Hasil dan Pembahasan.....	79
BAB V	SIMPULAN DAN SARAN.....	83
5.1	Simpulan.....	83
5.2	Saran.....	84
DAFTAR PUSTAKA		86
LAMPIRAN		90



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	8
Tabel 2. 2 Perbandingan Metode	25
Tabel 3. 1 Perbandingan Metode Data Mining [30].....	29



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

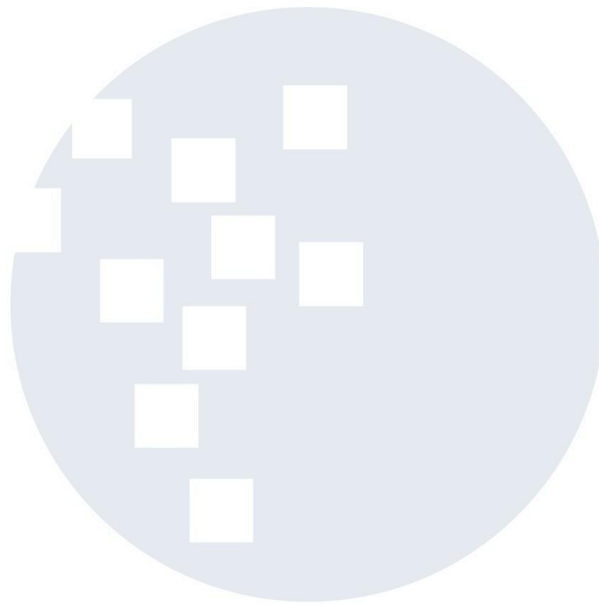
Gambar 2. 1 Gambaran Model SVM (spotintelligence.com)	17
Gambar 2. 2 Gambaran Model Naive Bayes (pemrogramanmatlab.com).....	18
Gambar 2. 3 Gambaran Model LSTM (medium.com)	20
Gambar 2. 4 Tahapan CRISP-DM	22
Gambar 3. 1 Diagram Alur Penelitian	28
Gambar 4. 1 Install & Import Library	37
Gambar 4. 2 Scraping data & membuat dataframe baru.....	38
Gambar 4. 3 Hasil Scraping Data.....	39
Gambar 4. 4 Hasil Purposive Sampling	39
Gambar 4. 5 Membaca data yang sudah terbentuk	40
Gambar 4. 6 Filter data dan mengganti nama fitur	40
Gambar 4. 7 Informasi Dataframe yang digunakan	41
Gambar 4. 8 Drop Duplicate Data.....	41
Gambar 4. 9 Cleaning Data.....	42
Gambar 4. 10 Case Folding.....	43
Gambar 4. 11 Tokenization.....	44
Gambar 4. 12 Proses Stopword Removal	44
Gambar 4. 13 Hasil Stopword Removal	44
Gambar 4. 14 Import kebutuhan untuk Stemming Data	45
Gambar 4. 15 Proses dan Hasil Stemming Data	45
Gambar 4. 16 Proses menyimpan data hasil preprocessing	46
Gambar 4. 17 Membaca data hasil preprocessing.....	46
Gambar 4. 18 Validasi data hasil preprocessing	47
Gambar 4. 19 Menghapus data kosong.....	47
Gambar 4. 20 Labelling dengan Lexicon Based	48
Gambar 4. 21 Dataframe dengan hasil labelling.....	48
Gambar 4. 22 Grafik distribusi sentimen	49
Gambar 4. 23 Menyimpan data hasil labelling	49
Gambar 4. 24 Import libraries untuk TF-IDF	50
Gambar 4. 25 Proses TF-IDF dan Split Data (Before SMOTE).....	51
Gambar 4. 26 Proses TF-IDF dan Split Data (<i>SMOTE</i>).....	52
Gambar 4. 27 Proses Label Encoding.....	52
Gambar 4. 28 Proses Tokenisasi dan Padding untuk LSTM before SMOTE.....	54
Gambar 4. 29 Proses Tokenisasi dan Padding untuk LSTM SMOTE	55
Gambar 4. 30 Word Cloud hasil Data Preparation	56
Gambar 4. 31 Model SVM Before SMOTE	57
Gambar 4. 32 Model SVM <i>After SMOTE</i>	59
Gambar 4. 33 Model Naive Bayes Before SMOTE.....	60
Gambar 4. 34 Model Naive Bayes SMOTE	62
Gambar 4. 35 Model LSTM Before SMOTE	63

Gambar 4. 36 Model LSTM SMOTE	64
Gambar 4. 37 Evaluasi Model SVM.....	66
Gambar 4. 38 Evaluasi Model SVM Confusion Matrix	67
Gambar 4. 39 Evaluasi Model SVM After SMOTE.....	68
Gambar 4. 40 Evaluasi Model Naive Bayes Before SMOTE.....	69
Gambar 4. 41 Confusion Matrix Naive Bayes Before SMOTE.....	71
Gambar 4. 42 Hasil Evaluasi Naive Bayes After SMOTE	72
Gambar 4. 43 Evaluasi Model LSTM before SMOTE	73
Gambar 4. 44 Confusion Matrix Model LSTM before <i>SMOTE</i>	74
Gambar 4. 45 Hasil Evaluasi Model LSTM After <i>SMOTE</i>	75
Gambar 4. 46 Program sederhana Analisis Sentimen berbasis SVM	77
Gambar 4. 47 Contoh Case Analisis Sentimen	78
Gambar 4. 48 Contoh Case Analisis Sentimen	79
Gambar 4. 49 Tabel Perbandingan Evaluasi	80
Gambar 4. 50 Model dengan akurasi tertinggi.....	81



DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	90
Lampiran B Form Konsultasi Bimbingan.....	91



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA