



**Analisis Sentimen Wacana Redenominasi Rupiah dengan
Lexicon-Based dan Perbandingan SVM, Random Forest,
Logistic Regression**

**TUGAS AKHIR
SKRIPSI**

YUDHA ANDIKA ISTANTO
41522010032

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**Analisis Sentimen Wacana Redenominasi Rupiah dengan
Lexicon-Based dan Perbandingan SVM, Random Forest,
Logistic Regression**

**TUGAS AKHIR
SKRIPSI**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

YUDHA ANDIKA ISTANTO

41522010032

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Yudha Andika Istanto

NIM : 41522010032

Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

“Analisis Sentimen Wacana Redenominasi Rupiah dengan Lexicon-Based dan Perbandingan SVM, Random Forest, Logistic Regression” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

UNIVERSITAS
MERCU BUANA

Jakarta, 25 Desember 2025



Yudha Andika Istanto

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
The undersigned, hereby declare that the scientific work written by

Nama /Name	: Yudha Andika Istanto
NIM /Student ID Number	: 41522010032
Program Studi /Study Program	: Teknik Informatika

Dengan Judul Tugas Akhir

The title:

“Analisis Sentimen Wacana Redenominasi Rupiah dengan Lexicon-Based dan Perbandingan SVM, Random Forest, Logistic Regression”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

Similarity checks have been carried out with the Turnitin system on the date:

5 Januari 2026

dengan nilai persentase sebesar :

with a percentage value of:

30%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. *declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek *similarity* turnitin:

Turnitin similarity report file

https://drive.google.com/file/d/1EWvWZciqqxZ83EzRGdAb3rP1aWLd_JcU/view?usp=drive_link



Jakarta, 5 Januari 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto, S.Kom., B.Sc

NIK : 322970503

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : YUDHA ANDIKA ISTANTO
NIM : 41522010032
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir : Analisis Sentimen Wacana Redenominasi Rupiah
dengan Lexicon-Based dan Perbandingan SVM,
Random Forest, Logistic Regression

Telah berhasil dipertahankan pada sidang tanggal 20 Januari 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing



Fauzi Nur Iman, S.Kom, M.Kom.

NIDN/NUPTK: 0318088903

MERCU BUANA

Jakarta, 26 Januari 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jekonowo, S.Si., MTI
NIDN/NUPTK: 0320037002

Ketua Program Studi
Teknik Informatika



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Bapak Fauzi Nur Iman, S.Kom, M.Kom selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mendukung saya selama berkuliah.
6. Sri Putri Marshanda, yang dengan penuh kesabaran senantiasa memberikan motivasi, dorongan moral, serta dukungan dalam berbagai proses yang saya jalani, khususnya selama penyusunan Tugas Akhir ini.
7. Seluruh teman-teman saya yang telah memberikan dukungan, semangat, serta kebersamaan yang berarti selama masa studi.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 11 Desember 2025

Yudha Andika Istanto

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB**

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

Nama : Yudha Andika Istanto
NIM : 41522010032
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Wacana Redenominasi Rupiah dengan
Lexicon-Based dan Perbandingan SVM, Random Forest,
Logistic Regression

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive 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 Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Desember 2025

Yang menyatakan,



Yudha Andika Istanto

**Analisis Sentimen Wacana Redenominasi Rupiah dengan Lexicon-Based dan
Perbandingan SVM, Random Forest,
Logistic Regression**

YUDHA ANDIKA ISTANTO

ABSTRAK

Penelitian ini menganalisis sentimen masyarakat terhadap wacana redenominasi rupiah 2025 di Twitter menggunakan pendekatan *lexicon-based* InSet Lexicon dan perbandingan algoritma *machine learning* (SVM, *Random Forest*, *Logistic Regression*). Data dikumpulkan melalui *web scraping* menghasilkan 3.559 *tweet*, setelah deduplikasi diperoleh 1.982 *tweet* unik. *Preprocessing* dilakukan melalui enam tahapan: *cleaning*, *case folding*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming* Sastrawi. Pelabelan *InSet Lexicon* menghasilkan distribusi sentimen positif 794 *tweet* (40,06%), negatif 630 *tweet* (31,79%), dan netral 558 *tweet* (28,15%), menunjukkan masyarakat cenderung optimis terhadap kebijakan ini. Ekstraksi fitur TF-IDF menghasilkan 1.000 fitur unik. Ketidakseimbangan kelas diatasi menggunakan SMOTE. Hasil evaluasi menunjukkan SMOTE meningkatkan performa semua model: SVM mencapai akurasi tertinggi 75% (peningkatan +7%), mengungguli *Random Forest* (72%, +5%) dan *Logistic Regression* (71%, +2%). SVM terbukti sebagai model terbaik dengan *macro average* F1-score 0,75 dan performa seimbang pada seluruh kelas. Penelitian ini berkontribusi dalam pemetaan opini publik berbasis data media sosial dan membuktikan efektivitas kombinasi *lexicon-based* dan *machine learning* dengan penanganan *imbalanced data* untuk analisis sentimen bahasa Indonesia.

Kata kunci: Analisis sentimen, redenominasi rupiah, SVM, *random forest*, *logistic regression*

**Analisis Sentimen Wacana Redenominasi Rupiah dengan Lexicon-Based dan
Perbandingan SVM, Random Forest,
Logistic Regression**

s

YUDHA ANDIKA ISTANTO

ABSTRACT

This study analyzes public sentiment toward the discourse on the 2025 rupiah redenomination on Twitter using a lexicon-based approach with the InSet Lexicon and a comparison of machine learning algorithms (Support Vector Machine, Random Forest, and Logistic Regression). Data were collected through web scraping, yielding 3,559 tweets, and after deduplication, 1,982 unique tweets were obtained. Preprocessing was conducted through six stages: cleaning, case folding, normalization, tokenization, stopword removal, and stemming using Sastrawi. Sentiment labeling based on the InSet Lexicon produced a distribution of 794 positive tweets (40.06%), 630 negative tweets (31.79%), and 558 neutral tweets (28.15%), indicating that public sentiment tends to be optimistic toward this policy. Feature extraction using TF-IDF generated 1,000 unique features. Class imbalance was addressed using the Synthetic Minority Over-sampling Technique (SMOTE). The evaluation results show that SMOTE improved the performance of all models: SVM achieved the highest accuracy of 75% (an increase of 7%), outperforming Random Forest (72%, +5%) and Logistic Regression (71%, +2%). SVM was proven to be the best-performing model, achieving a macro-average F1-score of 0.75 and balanced performance across all sentiment classes. This study contributes to the mapping of public opinion based on social media data and demonstrates the effectiveness of combining lexicon-based methods and machine learning with imbalanced data handling for Indonesian sentiment analysis.

Keywords: *Sentiment analysis, rupiah redenomination, SVM, random forest, logistic regression*

DAFTAR ISI

HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.4.1 Manfaat Teoritis.....	4
1.4.2 Manfaat Praktis	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu	6
2.2 Teori Utama	14
2.2.1 Media Sosial Twitter sebagai Sumber Data Analisis Sentimen.....	14
2.2.2 Web Scraping untuk Pengumpulan Data Twitter	15
2.2.3 Analisis Sentimen (Sentiment Analysis)	15
2.2.4 Text Preprocessing dengan Natural Language Toolkit (NLTK)	16
2.2.5 Google Colaboratory sebagai Platform Komputasi	16
2.2.6 Redenominasi Rupiah sebagai Kebijakan Moneter	17
2.2.7 Pendekatan Machine Learning untuk Klasifikasi Sentimen	17

2.2.8 Support Vector Machine (SVM).....	18
2.2.9 Random Forest	19
2.2.10 Logistic Regression.....	20
2.2.11 Feature Extraction dengan TF-IDF	21
2.2.12 Synthetic Minority Over-sampling Technique (SMOTE)	25
2.2.13 Confusion Matrix	26
2.2.14 Evaluation Metrics	28
2.3 Teori Pendukung	30
2.3.1 Media Sosial Twitter sebagai Sumber Data Analisis Sentimen.....	30
2.3.1 Analisis Sentimen (<i>Sentiment Analysis</i>)	30
2.3.3 Ekstraksi Fitur dengan TF-IDF	31
2.3.4 Algoritma Klasifikasi.....	32
2.3.5 Evaluasi Kinerja Model (<i>Model Evaluation</i>).....	33
2.3.6 Pendekatan Lexicon-Based untuk Analisis Sentimen.....	33
2.3.7 Knowledge Discovery in Database (KDD).....	33
BAB III METODE PENELITIAN	35
3.1 Jenis Penelitian	35
3.2 Tahapan Penelitian	35
3.2.1 Pengumpulan Data	36
3.2.2 Eksplorasi Data	37
3.2.3 Prapemrosesan Data.....	37
3.2.4 Pelabelan Sentimen (<i>Labeling</i>)	39
3.2.5 Ekstraksi Fitur (TF-IDF).....	39
3.2.6 Pembagian Data <i>Training</i> dan <i>Testing</i>	40
3.2.7 Penanganan Ketidakseimbangan Data (SMOTE).....	40
3.2.8 Penerapan Model Algoritma	40
3.2.9 Evaluasi Model	42
BAB IV PEMBAHASAN.....	43
4.1 Pengumpulan Data	43
4.1.1 Proses Deduplikasi Data	43
4.2 Preprocessing Data.....	44
4.2.1 Hasil Cleaning Data	44
4.2.2 Hasil Case Folding dan Normalisasi.....	46
4.2.3 Hasil Tokenizing dan Stopword Removal	49
4.2.4 Hasil Stemming.....	51
4.3 Hasil Pelabelan Sentimen dengan Lexicon-Based.....	53
4.3.1 Distribusi Hasil Pelabelan Sentimen.....	54
4.3.2 Analisis Kata Dominan per Sentimen.....	57
4.4 Hasil Feature Extraction dengan TF-IDF.....	59
4.5 Hasil Modeling Before SMOTE	62
4.5.1 Classification Report Model Before SMOTE.....	63

4.5.2 Perbandingan Akurasi Model Before SMOTE	67
4.6 Penerapan SMOTE untuk Handling Imbalance Data	67
4.6.1 Perubahan Distribusi Data Before vs After SMOTE	69
4.7 Hasil Modeling After SMOTE.....	71
4.7.2 Classification Report Model After SMOTE	73
BAB V KESIMPULAN DAN SARAN	78
5.1 Kesimpulan	78
5.2 Saran.....	78
DAFTAR PUSTAKA.....	80
LAMPIRAN.....	83



DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	6
Tabel 2. 2 Struktur Confusion Matrix untuk Multi-class Classification	27
Tabel 4. 1 Hasil Proses Deduplikasi Data	44
Tabel 4. 2 Classification Report Logistic Regression Before SMOTE	63
Tabel 4. 3 Classification Report Random Forest Before SMOTE	64
Tabel 4. 4 Classification Report SVM Before	65
Tabel 4. 5 Akurasi Model Before SMOTE	67
Tabel 4. 6 Akurasi Model After SMOTE	71
Tabel 4. 7 Classification Report Logistic Regression After SMOTE	73
Tabel 4. 8 Classification Report Random Forest After SMOTE	74
Tabel 4. 9 Classification Report SVM After SMOTE	76



DAFTAR GAMBAR

Gambar 3. 1 Diagram Alur Metodologi Penelitian.....	36
Gambar 4. 1 Code Cleaning (Pembersihan Teks).....	46
Gambar 4. 2 Hasil Cleaning Data	46
Gambar 4. 3 Code Case Folding	48
Gambar 4. 4 Code Normalisasi	48
Gambar 4. 5 Hasil Case Folding dan Normalisasi	49
Gambar 4. 6 Code Tokenizing	50
Gambar 4. 7 Code Stopword Removal	51
Gambar 4. 8 Hasil Tokenizing dan Stopword Removal	51
Gambar 4. 9 Code Stemming	53
Gambar 4. 10 Hasil Stemming	53
Gambar 4. 11 Pelabelan Data dengan Lexicon - Based	56
Gambar 4. 12 Distribusi Sentimen Hasil Pelabelan	56
Gambar 4. 13 WordCloud Positif.....	58
Gambar 4. 14 WordCloud Negatif	58
Gambar 4. 15 WordCloud Netral	59
Gambar 4. 16 Feature Extraction TF-IDF	62
Gambar 4. 17 Confusion Matrix Logistic Regression Before SMOTE	63
Gambar 4. 18 Confusion Matrix Random Forest Before SMOTE	64
Gambar 4. 19 Confusion Matrix SVM Before SMOTE	65
Gambar 4. 20 Klasifikasi Before SMOTE	66
Gambar 4. 21 Visualisasi Akurasi Model Before SMOTE.....	67
Gambar 4. 22 Code Penerapan SMOTE	68
Gambar 4. 23 Code Jumlah Analisis Sentimen Before SMOTE	69
Gambar 4. 24 Distribusi Data Before SMOTE	69
Gambar 4. 25 Code Jumlah Analisis Sentimen After SMOTE.....	70
Gambar 4. 26 Distribusi Data After SMOTE	70
Gambar 4. 27 Visualisasi Perbandingan Akurasi Before vs After SMOTE	72
Gambar 4. 28 Confusion Matrix Logistic Regression After SMOTE	73
Gambar 4. 29 Confusion Matrix Random Forest After SMOTE	75
Gambar 4. 30 Confusion Matrix SVM After SMOTE.....	76

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi	83
Lampiran 2. Curriculum Vitae	84
Lampiran 3. Surat Pernyataan & Pengalihan HAKI	85
Lampiran 4. Sertifikat BNSP	87
Lampiran 5. Form Revisi Dosen Penguji.....	88

