

**IMPLEMENTASI ALGORITMA SUPPORT VECTOR
MACHINE UNTUK KLASIFIKASI OPINI PUBLIK
TERHADAP INSIDEN PERETASAN INDODAX**



SKRIPSI

**RAFLES KRISTIYANTO
00000032818**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**IMPLEMENTASI ALGORITMA SUPPORT VECTOR
MACHINE UNTUK KLASIFIKASI OPINI PUBLIK
TERHADAP INSIDEN PERETASAN INDODAX**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

RAFLES KRISTIYANTO
00000032818

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Raffles Kristiyanto
Nomor Induk Mahasiswa : 00000032818
Program Studi : Informatika

Skripsi dengan judul:

Implementasi Algoritma Support Vector Machine untuk Klasifikasi Opini Publik terhadap Insiden Peretasan Indodax

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, 03 Juli 2025



(Raffles Kristiyanto)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI OPINI PUBLIK TERHADAP INSIDEN PERETASAN INDODAX

oleh

Nama : Rafles Kristiyanto
NIM : 00000032818
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Rabu, 09 Juli 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Yustinus Widya Wiratama, S.Kom.,
M.Sc.)

NIDN: 0305128901

Penguji



(Angga Aditya Permana, S.Kom.,
M.Kom.)

NIDN: 0407128901

Pembimbing



(Dr. Maria Irmina Prasetyowati, S.Kom., M.T.)

NIDN: 0725057201

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Rafles Kristiyanto
NIM : 00000032818
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Implementasi Algoritma Support
Vector Machine untuk Klasifikasi
Opini Publik terhadap Insiden
Peretasan Indodax

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

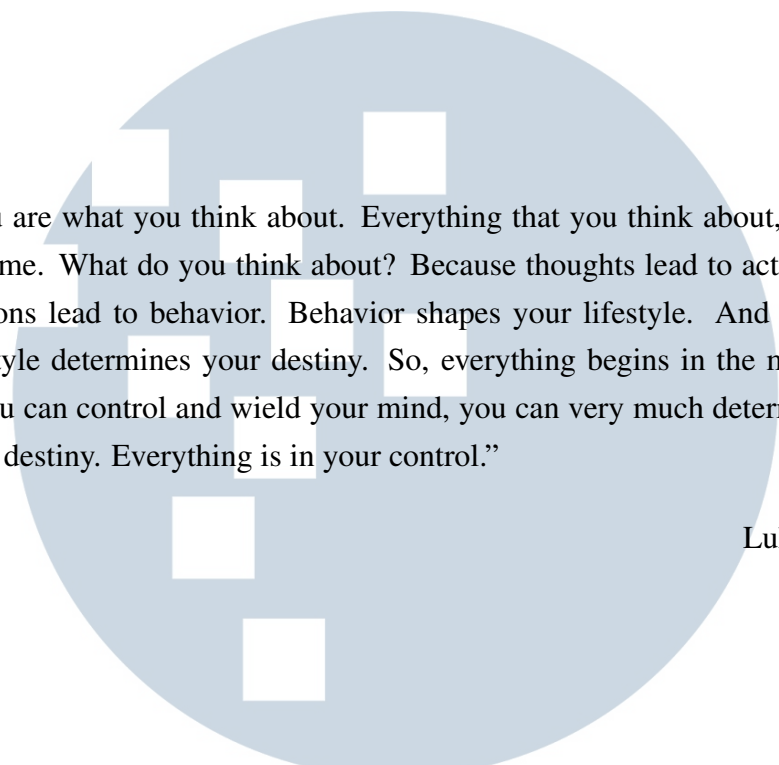
- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang 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
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 03 Juli 2025

Yang menyatakan


Rafles Kristiyanto

HALAMAN PERSEMBAHAN / MOTTO



"You are what you think about. Everything that you think about, you become. What do you think about? Because thoughts lead to actions. Actions lead to behavior. Behavior shapes your lifestyle. And your lifestyle determines your destiny. So, everything begins in the mind. If you can control and wield your mind, you can very much determine your destiny. Everything is in your control."

Luke Belmar

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan rahmat-Nya sehingga skripsi yang berjudul *Implementasi Algoritma Support Vector Machine untuk Klasifikasi Opini Publik Terhadap Insiden Peretasan Indodax* ini dapat diselesaikan dengan baik sebagai salah satu syarat memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Saya menyadari bahwa penyusunan skripsi ini tidak terlepas dari dukungan, bimbingan, dan bantuan berbagai pihak. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Dr. Maria Irminda Prasetyowati, S.Kom., M.T., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan skripsi ini.
6. Kepada teman-teman di kampus maupun di luar kampus, khususnya Tasya, Verel, Leonardo, Janu, David, Giovanni, dan Juan, serta teman-teman lainnya yang tidak dapat disebutkan satu per satu, yang telah memberikan semangat, ide, dukungan, doa, dan motivasi selama proses penyusunan skripsi ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 03 Juli 2025



Rafles Kristiyanto

IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI OPINI PUBLIK TERHADAP INSIDEN PERETASAN

INDODAX

Raffles Kristiyanto

ABSTRAK

Seiring meningkatnya penggunaan *cryptocurrency* di Indonesia, isu keamanan siber pada *platform* pertukaran aset digital seperti Indodax menjadi perhatian serius. Penelitian ini mengimplementasikan algoritma *Support Vector Machine* (*SVM*) untuk menganalisis sentimen publik terhadap insiden peretasan Indodax pada September 2024. Data dikumpulkan dari media sosial dan diproses melalui tahapan *preprocessing*, seperti pembersihan teks, *stemming*, dan ekstraksi fitur menggunakan *TF-IDF*. Model dilatih dengan *GridSearchCV* untuk optimasi parameter, diuji dengan berbagai konfigurasi *N-gram*, serta menggunakan *SMOTE* untuk menangani ketidakseimbangan kelas. Model terbaik, yaitu *Linear SVM* dengan konfigurasi *TF-IDF N-gram* (1,2) dan *SMOTE*, mencapai akurasi 78,73% dan macro average F1-score 76,14%. Hasil klasifikasi terhadap data baru menunjukkan bahwa opini publik didominasi sentimen negatif. Penelitian ini memberikan kontribusi dalam memahami persepsi masyarakat terhadap insiden siber dan sebagai masukan bagi platform kripto dalam mitigasi dan strategi komunikasi.

Kata kunci: Analisis sentimen, Indodax, keamanan siber, opini publik, Support Vector Machine, TF-IDF.

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

**IMPLEMENTATION OF THE SUPPORT VECTOR MACHINE ALGORITHM
FOR PUBLIC OPINION CLASSIFICATION ON THE INDODAX HACKING
INCIDENT**

Raffles Kristiyanto

ABSTRACT

As cryptocurrency adoption increases in Indonesia, cybersecurity on digital asset exchange platforms such as Indodax has become a critical concern. This study implements the Support Vector Machine (SVM) algorithm to analyze public sentiment toward the Indodax hacking incident in September 2024. Social media data were collected and processed through text cleaning, stemming, and feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF). The model was trained using GridSearchCV for parameter optimization, tested with various N-gram configurations, and applied the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. The best-performing model, Linear SVM with TF-IDF N-gram (1,2) and SMOTE, achieved 78.73% accuracy and a macro average F1-score of 76.14%. Classification of post-incident data indicated a predominance of negative sentiment. This study contributes to understanding public perceptions of cybersecurity incidents and offers strategic insights for crypto platform providers in crisis response, risk mitigation, and communication strategies following security breaches.

Keywords: Cybersecurity, Indodax, Public opinion, Sentiment analysis, Support Vector Machine, TF-IDF

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	6
2.1 Analisis Sentimen	6
2.2 Text Mining	6
2.3 Text Preparation	6
2.4 Term Frequency-Inverse Document Frequency (TF-IDF)	7
2.5 Train Test Split	8
2.6 Support Vector Machine	8
2.7 Confusion Matrix	10
BAB 3 METODOLOGI PENELITIAN	12
3.1 Alur Penelitian	12
3.1.1 Studi Literatur	12
3.1.2 Pengumpulan Data	12
3.1.3 Perancangan Sistem	13
3.1.4 Pengujian	13
3.1.5 Evaluasi	13
3.2 Gambaran Umum Sistem	13
3.2.1 Scraping Data	14
3.2.2 Data Preprocessing	15
3.2.3 Term Frequency Inverse Document Frequency (TF-IDF)	16
3.2.4 Splitting Data	17
3.2.5 Modeling Support Vector Machine	17
BAB 4 HASIL DAN DISKUSI	19
4.1 Spesifikasi Sistem	19
4.1.1 Spesifikasi Perangkat Keras	19
4.1.2 Spesifikasi Perangkat Lunak	19
4.2 Analisis Data	19
4.2.1 Import Library	20

4.2.2	Pengumpulan Data	21
4.2.3	WordCloud Data Mentah	22
4.2.4	WordCloud Data Bersih	23
4.3	Preprocessing Data	23
4.3.1	Cleaning Data	24
4.3.2	Case Folding	25
4.3.3	Text Normalization	26
4.3.4	Text Tokenizing	28
4.3.5	Stopword	29
4.3.6	Stemming	30
4.3.7	Labeling Data	31
4.3.8	Ekstraksi Fitur Menggunakan TF-IDF	34
4.3.9	Train-Test Split Data dan Model Suport Vector Machine	35
4.3.10	Evaluasi Model	36
4.3.11	Visualisasi Learning Curve dan Validation Curve	37
4.4	Pengujian Model	39
4.4.1	Uji Coba 1: Tuning Parameter Model SVM	40
4.4.2	Uji Coba 2: Variasi <i>N-Gram</i> pada Model SVM	42
4.4.3	Uji Coba 3: Penanganan Ketidakseimbangan Data dengan <i>SMOTE</i>	44
4.5	Pelabelan Sentimen pada Data Baru	46
4.5.1	Evaluasi	46
4.5.2	Validation Curve Model Final	47
4.5.3	Prediksi Teks	47
BAB 5	SIMPULAN DAN SARAN	49
5.1	Simpulan	49
5.2	Saran	49
	DAFTAR PUSTAKA	51



DAFTAR TABEL

Tabel 2.1	Confusion Matrix	10
Tabel 4.1	Cleaning Text	25
Tabel 4.2	Contoh Perubahan Teks setelah Case Folding	26
Tabel 4.3	Contoh Hasil Case Folding dan Normalisasi	28
Tabel 4.4	Contoh Hasil Tokenisasi setelah Normalisasi	29
Tabel 4.5	Contoh Hasil Penghapusan Stopword setelah Tokenisasi	30
Tabel 4.6	Contoh Hasil Stopword Removal dan Stemming	31
Tabel 4.7	Hasil evaluasi model SVM berdasarkan metrik klasifikasi	36
Tabel 4.8	Perbandingan hasil evaluasi Linear SVM dan RBF SVM	42
Tabel 4.9	Hasil evaluasi F1 Macro terhadap variasi N-gram	44
Tabel 4.10	Perbandingan performa model SVM sebelum dan sesudah SMOTE	45
Tabel 4.11	Evaluasi performa model final (Linear SVM + TF-IDF (1,2) + SMOTE)	46



DAFTAR GAMBAR

Gambar 1.1	Statistik bursa kripto yang paling banyak digunakan	1
Gambar 2.1	Ilustrasi <i>hyperplane</i> , margin, dan <i>support vectors</i> dalam SVM.	10
Gambar 3.1	Struktur alur penelitian	12
Gambar 3.2	Gambaran umum sistem	14
Gambar 3.3	Flowchart Scraping Data	15
Gambar 3.4	Flowchart Data Preprocessing	16
Gambar 3.5	Flowchart TF-IDF	17
Gambar 3.6	Flowchart Splitting Data	17
Gambar 3.7	Flowchart Modeling SVM	18
Gambar 4.1	Tampilan Data 5 Baris Pertama Hasil Pengumpulan	22
Gambar 4.3	WordCloud Data Bersih	23
Gambar 4.4	Potongan Code Inset Lexicon	31
Gambar 4.5	Hasil Labeling	33
Gambar 4.6	Pie Chart Distribusi Sentiment	34
Gambar 4.7	Visualisasi <i>Confusion Matrix</i>	37
Gambar 4.8	Learning Curve model SVM berdasarkan F1 macro score	38
Gambar 4.9	Validation Curve model SVM terhadap variasi nilai parameter <i>C</i>	39
Gambar 4.10	Validation Curve pada model Linear SVM (TF-IDF (1,2) + SMOTE)	47
Gambar 4.11	Hasil prediksi data baru beserta fitur yang paling berpengaruh terhadap klasifikasi	48



DAFTAR KODE

Kode 4.1	Import Library	20
Kode 4.2	Pengambilan dan Pembersihan Data	21
Kode 4.3	Cleaning Data	24
Kode 4.4	Fungsi dan Penerapan Case Folding	25
Kode 4.5	Fungsi dan Penerapan Normalisasi Text	26
Kode 4.6	Fungsi dan Penerapan Tokenizing	28
Kode 4.7	Fungsi dan Penerapan Stopword Removal	29
Kode 4.8	Fungsi dan Penerapan Stemming	30
Kode 4.9	Potongan kode fungsi penentuan label	32
Kode 4.10	Potongan kode transformasi teks menjadi TF-IDF	34
Kode 4.11	Potongan kode Train-Test Split	35
Kode 4.12	Potongan kode implementasi uji coba GridSearchCV	40
Kode 4.13	Potongan kode implementasi uji coba variasi N-Gram	42
Kode 4.14	Potongan kode implementasi uji coba SMOTE	44



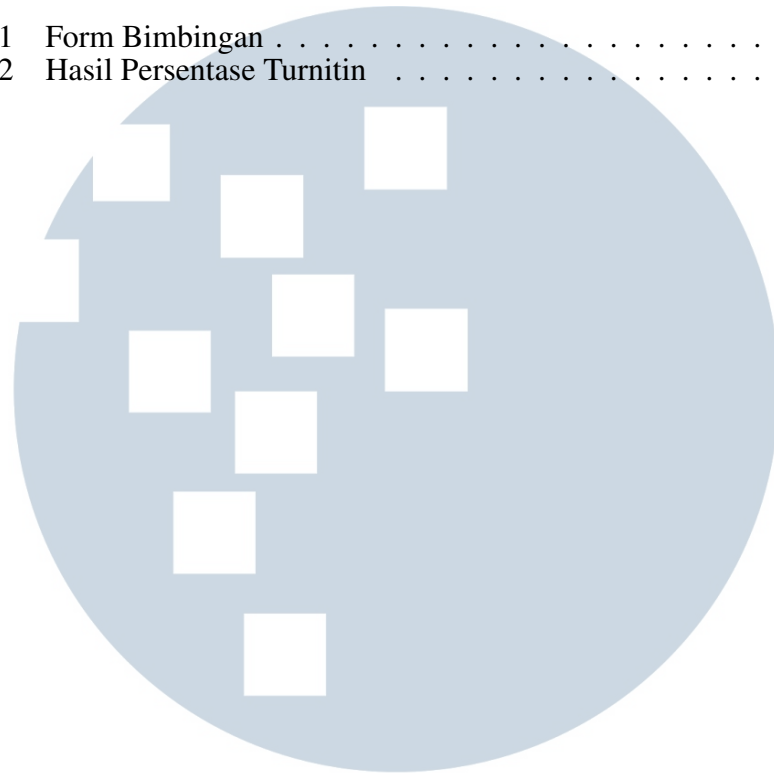
DAFTAR RUMUS

Rumus 2.1	<i>Term Frequency (TF)</i>	7
Rumus 2.2	<i>Inverse Document Frequency (IDF)</i>	7
Rumus 2.3	<i>Term Frequency - Inverse Document Frequency (IDF)</i>	7
Rumus 2.4	<i>Hyperplane - SVM</i>	9
Rumus 2.5	<i>Margin - SVM</i>	9
Rumus 2.6	<i>Accuracy</i>	11
Rumus 2.7	<i>Recall</i>	11
Rumus 2.8	<i>Precision</i>	11
Rumus 2.9	<i>F1-Score</i>	11



DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	55
Lampiran 2	Hasil Persentase Turnitin	57



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA