

SKRIPSI
PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE PADA
ANALISIS SENTIMEN TERHADAP PELUNCURAN DANANTARA



PETRA ALDEVAND HOSYO

NIM : 215610041

PROGRAM STUDI SISTEM INFORMASI
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2025

SKRIPSI

PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP PELUNCURAN DANANTARA

Diajukan sebagai salah satu syarat untuk menyelesaikan studi pada

**Program Sarjana
Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia Yogyakarta**

Disusun Oleh

PETRA ALDEVAND HOSYO

NIM : 215610041

**PROGRAM STUDI
SISTEM INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR

Judul : Penerapan Algoritma Support Vector Machine Pada Analisis Sentimen Terhadap Peluncuran Danantara

Nama : Petra Aldevand Hosyo

NIM : 215610041

Program Studi : Sistem Informasi

Program : Sarjana

Semester : Genap

Tahun Akademik : 2024/2025

Telah diperiksa dan disetujui untuk diujikan
di hadapan Dewan Penguji Tugas Akhir



Yogyakarta, 24 Juli 2025

Dosen Pembimbing,

Cosmas Haryawan, S.TP., S. Kom., M.Cs..

NIDN: 0519067401

HALAMAN PENGESAHAN

PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE PADA

ANALISIS SENTIMEN TERHADAP PELUNCURAN DANANTARA

Telah dipertahankan di depan Dewan Penguji dan dinyatakan diterima untuk
memenuhi sebagian persyaratan guna memperoleh

Sarjana Komputer
Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia
Yogyakarta

Yogyakarta, 24 Juli 2025

Dewan Penguji	NIDN	Tandatangan
1. Deborah Kurniawati, S.Kom., M.Cs.	0511107301	
2. Cosmas Haryawan, S.TP., S. Kom., M.Cs.	0519067401	
3. Dr. Asyahri Hadi Nasyuha, S.Kom., M.Kom.	0129048601	

Mengetahui

Ketua Program Studi Program Studi



Deborah Kurniawati, S.Kom., M.Cs.

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa naskah Tugas Akhir ini belum pernah diajukan untuk memperoleh gelar Nama Gelar Sarjana di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 24 Juli 2025



Petra Aldevand Hosyo

NIM: 215610041

HALAMAN PERSEMBAHAN

Dengan rasa syukur dan hormat yang mendalam, karya ini penulis persembahkan
kepada:

Ayahanda Oktovianus Hosyo,
yang selalu menjadi teladan dan sumber semangat dalam setiap langkah hidup.

Ibunda Yulianti Kambu,
atas kasih sayang, doa, dan dukungan yang tak pernah henti sepanjang perjalanan
ini.

Yulani Yesimel Nauw,
yang senantiasa hadir memberi semangat dan motivasi dalam setiap proses yang
dijalani.

Dan untuk diri sendiri,
*"Last but not least, I wanna thank me, I wanna thank me for believing in me,
I wanna thank me for doing all this hard work,
I wanna thank me for having no days off,
I wanna thank me for... for never quitting,
I wanna thank me for always being a giver and try give more than I receive,
I wanna thank me for try and do more right than wrong,
I wanna thank me for just being me at all times."*

(Calvin Cordozar Broadus Jr.)

PRAKATA

Puji syukur kehadiran Tuhan Yang Maha Esa atas rahmat-Nya, sehingga penelitian berjudul PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP PELUNCURAN DANANTARA dapat diselesaikan. Penelitian ini bertujuan untuk mengklasifikasikan dan menganalisis sentimen publik terhadap peluncuran lembaga tersebut.

Penyelesaian penelitian ini dimungkinkan berkat dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, ucapan terima kasih disampaikan kepada:

1. Cosmas Haryawan, S.TP., S.Kom., M.CS., selaku dosen pembimbing yang telah meluangkan waktu, memberikan arahan, masukan, serta motivasi yang tak ternilai harganya sejak awal hingga akhir penelitian ini.
2. Bapak/Ibu Dosen di Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia yang telah membekali dengan ilmu pengetahuan dan wawasan selama masa studi.
3. Keluarga tercinta, Ayah Oktovianus Hosyo, Ibu Yulianti Kambu, atas doa, dukungan moral, dan kasih sayang yang tiada henti.
4. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan kontribusi, baik secara langsung maupun tidak langsung, dalam penyelesaian penelitian ini dan selama masa pendidikan.

DAFTAR ISI

SKRIPSI	i
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TUGAS AKHIR	v
HALAMAN PERSEMBAHAN	vi
PRAKATA	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
INTISARI	xii
ABSTRACT	xiii
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Ruang Lingkup	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II	6
TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori	8
2.2.1 Data Mining	9
2.2.2 Analisis Sentimen	10
2.2.3 X (Twitter)	11
2.2.4 Natural Language Processing (NLP)	12
2.2.5 Ekstraksi Fitur	13
2.2.6 Machine Learning	14
2.2.7 Pembagian Data	15

2.2.8 Penanganan Data Tidak Seimbang	15
2.2.9 Support Vector Machine	16
2.2.10 Pelabelan	18
2.2.11 Evaluasi Hasil.....	21
BAB III.....	22
METODOLOGI PENELITIAN.....	22
3.1 Bahan/Data.....	22
3.2 Peralatan.....	23
3.3 Prosedur dan Pengumpulan Data	24
3.4 Analisis dan Rancangan Sistem	26
BAB IV	29
IMPLEMENTASI DAN PEMBAHASAN	29
4.1 Pengumpulan dan Penggabungan Data	29
4.1.1 Pembersihan Data dan Seleksi Fitur.....	30
4.1.2 Pra-Pemrosesan Teks.....	31
4.1.3 Pelabelan Sentimen Berbasis Lexicon.....	34
4.1.4 Seleksi dan Pembagian Data	35
4.1.5 Ekstraksi Fitur	36
4.1.6 Penerapan SMOTE.....	37
4.1.7 Pelatihan.....	38
4.1.8 Pengujian dan Evaluasi.....	39
4.2 Pembahasan.....	42
BAB V.....	44
PENUTUP	44
5.1 Kesimpulan	44
5.2 Saran	45
DAFTAR PUSTAKA.....	46
LAMPIRAN	49

DAFTAR GAMBAR

GAMBAR 2.1 SVM LINEAR (CHOLISSODIN & SOEBROTO, 2021)	17
GAMBAR 3.1 FLOWCHART TAHAPAN PENELITIAN ANALISIS SENTIMEN	28
GAMBAR 4.2 VISUALISASI DISTRIBUSI LABEL	34
GAMBAR 4.4 CONTOH KATA TERATAS DARI KEDUA KELAS DARI TF-IDF.....	37
GAMBAR 4.5 VISUALISASI SEBELUM DAN SETELAH PENYEIMBANGAN	38
GAMBAR 4.6 CLASSIFICATION REPORT SEBELUM SMOTE.....	39
GAMBAR 4.7 CONFUSION MATRIX SEBELUM SMOTE	40
GAMBAR 4.8 CLASSIFICATION SETELAH SMOTE	41
GAMBAR 4.9 CONFUSION MATRIX SESUDAH SMOTE.....	41

DAFTAR TABEL

TABEL 2.1 PENELITIAN TERDAHULU.....	8
TABEL 4.1 INFORMASI DATASET.....	30
TABEL 4.2 REKAP PROSES PEMBERSIHAN.....	31
TABEL 4.3 BEBERAPA BARIS KOLOM FULL_TEXT.....	31
TABEL 4.4 PROSES TRANSFORMASI TEKS.....	33
TABEL 4.5 REKAP DATA SETELAH PRA-PEMROSESAN.....	33
TABEL 4.6 BEBERAPA BARIS HASIL PELABELAN.....	35
TABEL 4.7 DISTRIBUSI DATA HASIL PEMBAGIAN.....	35
TABEL 4.8 HASIL PENERAPAN SMOTE.....	38

INTISARI

Lembaga Badan Pengelola Investasi Daya Anagata Nusantara (BPI Danantara) yang diluncurkan pada Senin, 24 Februari 2025 telah menimbulkan perbincangan serta pro-kontra di berbagai lapisan masyarakat. Penelitian ini bertujuan untuk melakukan klasifikasi serta analisis kecenderungan sentimen publik terhadap peluncuran lembaga tersebut dengan menggunakan algoritma *machine learning Support Vector Machine*. Data yang diperoleh berasal dari platform X (Twitter) yang dikumpulkan dengan bantuan *tools tweet-harvest* pada rentang waktu 24 Februari – 14 Maret 2025. Proses penelitian ini dimulai dengan pengumpulan data lalu seleksi fitur yang relevan yakni *full_text* kemudian data teks dinormalisasi dengan beberapa tahapan seperti *cleaning*, *case folding*, *cleansing* (URL, *mention*, *hashtag*, angka, tanda baca), tokenisasi, *stopword removal*, *stemming*, *Joining*, dan hapus kosong. Kemudian proses pemberian label dilakukan dengan pendekatan berbasis *lexicon* (lexicon based) dengan kamus InSet Lexicon, kemudian label netral dipisahkan. Data final (positif dan negatif) dibagi menjadi dua bagian dengan rasio 80:20. Proses ekstraksi fitur dilakukan dengan metode TF-IDF untuk mendapatkan representasi kata unik (fitur) lalu untuk mengatasi ketidakseimbangan yang terjadi pada data, teknik *oversampling* dengan SMOTE digunakan untuk menyeimbangkan sebaran data pada data latih. Proses klasifikasi dilakukan *Support Vector Machine* dengan *kernel* linear. Model yang dibangun memiliki performa yang cukup baik dengan akurasi 81.86%, dengan FI-score tertinggi pada kelas positif 87% dan untuk kelas negatif hanya mencapai 68%, artinya meskipun performa yang baik dalam mengklasifikasikan sentimen positif namun performanya sedikit kurang optimal untuk kelas negatif. Secara keseluruhan mayoritas opini publik terhadap peluncuran Danantara cenderung positif.

Kata Kunci: Analisis Sentimen, Support Vector Machine, Danantara

ABSTRACT

The Nusantara Anagata Investment Management Agency (BPI Danantara), launched on Monday, February 24, 2025, has sparked discussions and created both support and opposition among various segments of society. This research aims to classify and analyze the public sentiment trends towards the launch of this institution using the Support Vector Machine (SVM) machine learning algorithm. The data was collected from the X (Twitter) platform with the aid of the tweet-harvest tool, covering the period from February 24 to March 14, 2025. The research process began with data collection, followed by the selection of the relevant feature, namely `full_text`. The text data was then normalized through several stages, including cleaning, case folding, cleansing (removing URLs, mentions, hashtags, numbers, punctuation), tokenization, stopwords removal, stemming, joining, and removing empty entries. Subsequently, labeling was performed using a lexicon-based approach with the InSet Lexicon dictionary, and neutral labels were separated. The final data (positive and negative) was divided into two parts with an 80:20 ratio. Feature extraction was carried out using the TF-IDF method to obtain unique word representations (features). To address data imbalance, the SMOTE oversampling technique was applied to balance the data distribution in the training set. Classification was performed using Support Vector Machine with a linear kernel. The developed model demonstrated good performance with an accuracy of 81.86%, achieving the highest F1-score for the positive class at 87%, while the negative class reached only 68%. This indicates that although the model performs well in classifying positive sentiment, its performance is slightly less optimal for the negative class. Overall, the majority of public opinion regarding the launch of Danantara tends to be positive.

Keyword: Analisis Sentimen, Support Vector Machine, Danantara