

**ANALISIS PERFORMA SUPPORT VEKTOR MACHINE
UNTUK KLASIFIKASI TOPIK BERITA PADA MERAPI
UNCOVER**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

BRAVENDA KRISNA MURTI

21.11.3865

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS PERFORMA SUPPORT VEKTOR MACHINE
UNTUK KLASIFIKASI TOPIK BERITA PADA MERAPI
UNCOVER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

BRAVENDA KRISNA MURTI

21.11.3865

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS PERFORMA SUPPORT VEKTOR MACHINE UNTUK KLASIFIKASI TOPIK BERITA PADA MERAPI UNCOVER

yang disusun dan diajukan oleh

Nama Mahasiswa
BRAVENDA KRISNA MURTI

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 12 Desember 2025

Dosen Pembimbing,



Bayu Setiaji, M.kom.

NIK. 190302216

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS PERFORMA SUPPORT VEKTOR MACHINE UNTUK KLASIFIKASI TOPIK BERITA PADA MERAPI UNCOVER

yang disusun dan diajukan oleh

BRAVENDA KRISNA MURTI

21.11.3865

Telah dipertahankan di depan Dewan Penguji
pada tanggal 12 Desember 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ali Mustopa, S.Kom., M.Kom.

NIK. 190302192

Rumini, S.Kom., M.Kom.

NIK. 190302246

Bayu Setiaji, M.Kom.

NIK. 190302216

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 Desember 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : BRAVENDA KRISNA MURTI
NIM : 21.11.3865

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS PERFORMA SUPPORT VEKTOR MACHINE UNTUK
KLASIFIKASI TOPIK BERITA PADA MERAPI UNCOVER**

Dosen Pembimbing : Bayu Setiaji, M.kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 12 Desember 2025

Yang Menyatakan,



Bravenda krisna Murti

HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kepada Allah SWT atas segala limpahan rahmat dan ridhoNya yang telah memberikan kesehatan, kelancaran dan kekuatan. Atas segala karunia serta kemudahan yang Engkau berikan akhirnya skripsi ini dapat terselesaikan. Selain itu penulis juga berterimakasih kepada orang yang sangat berarti dalam pembuatan skripsi ini:

1. Kedua Orang Tua, Alm. Bapak Nur Isgiyanta dan Ibu Sri Yuliwanti, kakak Yugama Ciesna Murti, Istri tercinta Endang Sukesi Wulandari dan anak tersayang Priscilla Naeema Batari yang telah memberikan doa, motivasi, semangat yang tak ternilai harganya.
2. Dosen pembimbing, Bayu Setiaji, M.kom, yang telah membimbing dari awal hingga terselesaikannya skripsi.
3. Teman-teman dari Yogyakarta, Teman-teman Kontrakan Reborn, Teman-teman Universitas Amikom Yogyakarta (IF01 angkatan 2021), dan Teman-teman yang sudah banyak membantu saya dalam proses pengerjaan skripsi.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT atas karunia dan limpahan rahmat dan ridho-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul “ANALISIS PERFORMA SUPPORT VEKTOR MACHINE UNTUK KLASIFIKASI TOPIK BERITA PADA MERAPI UNCOVER”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Universitas Amikom Yogyakarta.

Dalam kesempatan ini, penulis menyampaikan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Kedua Orang Tua, Alm. Bapak Nur Isgiyanta dan Ibu Sri Yuliwanti, kakak Yugama Ciesna Murti, Istri tercinta Endang Sukesi Wulandari dan anak tersayang Priscilla Naeema Batari yang telah memberikan doa, motivasi, semangat yang tak ternilai harganya.
2. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas AMIKOM Yogyakarta.
3. Bapak Prof. Dr. Kusrini, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
4. Bapak Bayu Setiaji, M.kom, selaku dosen pembimbing penulis yang telah memberikan kritik, saran, waktu, motivasi dan bimbingan dalam skripsi ini sehingga skripsi ini dapat selesai dengan baik.
5. Teman-teman dari Yogyakarta, Teman-teman Kontrakan Reborn, Teman-teman Universitas Amikom Yogyakarta (IF01 angkatan 2021), dan Teman-teman yang sudah banyak membantu saya dalam proses pengerjaan skripsi.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, segala kritik dan saran yang membangun sangat diharapkan. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak.

Yogyakarta, 12 Desember 2025

Penulis

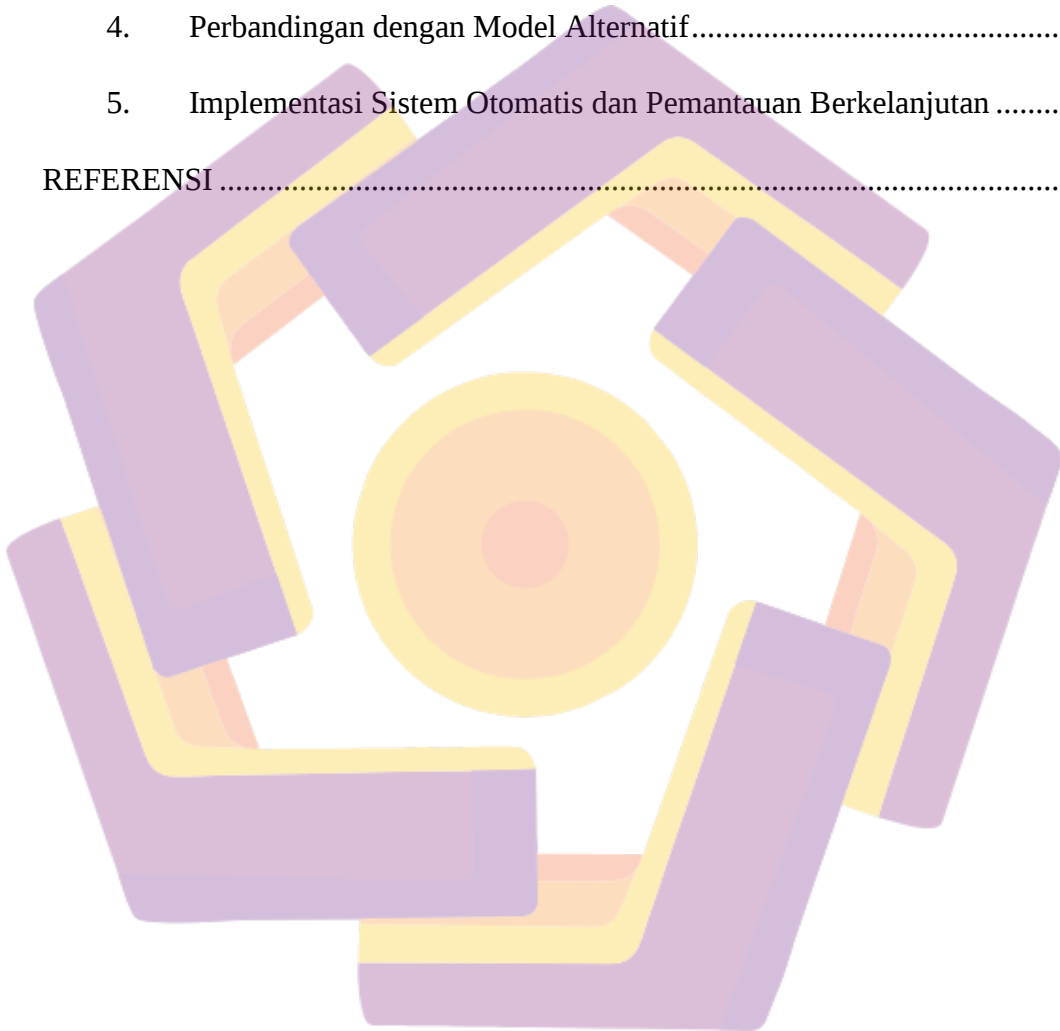
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xv
INTISARI	xvii
ABSTRACT.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	6

2.1	Studi Literatur	6
2.2	Dasar Teori	17
2.2.1.	Instagram.....	17
2.2.2.	Merapi Uncover	18
2.2.3.	Machine Learning	18
2.2.4.	Klasifikasi	19
2.2.5.	Klasifikasi Text.....	19
2.2.6.	Klasifikasi Berita Digital	20
2.2.7.	Preprocessing Text.....	20
2.2.8.	Term Frequency-Inverse Document Frequency (TF-IDF)	21
2.2.9.	Support Vector Machine (SVM).....	21
2.2.10.	Evaluasi Model Klasifikasi	26
BAB III METODE PENELITIAN		28
3.1	Objek Penelitian	28
3.2.	Alur Penelitian.....	29
1.	Pengumpulan Data.....	30
2.	Labeling Dataset	30
3.	Preprocessing Text.....	30
4.	Clean Dataset	30
5.	Data Splitting	30
6.	Term Frequency-Inverse Document Frequency (TF– DF).....	31
7.	Training Model Support Vector Machine (SVM)	31

8.	Evaluasi.....	35
3.3	Alat dan Bahan	36
3.3.1	Alat.....	36
3.3.2	Bahan	36
BAB IV HASIL DAN PEMBAHASAN		37
4.1	Pengumpulan Data	37
4.2	Hasil Labeling dan Pembentukan Kategori.....	38
4.3	Hasil Preprocessing dan Cleaning Teks.....	40
4.4	Pembagian Data Latih dan Uji	41
4.5	Representasi Fitur dengan TF-IDF	42
4.6	Hasil Pelatihan Model Support Vector Machine.....	42
4.7.	Hasil Evaluasi Model	43
1.	Ringkasan proses training.....	43
2.	Gambaran evaluasi umum	43
3.	Akumulasi Evaluasi Model per Kategori Topik.....	45
4.	Implikasi dari hasil evaluasi	48
BAB V PENUTUP		49
5.1	Kesimpulan.....	49
1.	Performa Keseluruhan Model.....	49
2.	Performa Per Kategori Topik.....	49
3.	Faktor Penentu Performa	50
4.	Kontribusi Penelitian	50

5.2	Saran.....	51
1.	Penyeimbangan Kelas dan Perluasan Data.....	51
2.	Optimasi Hyperparameter dan Eksplorasi Kernel	51
3.	Pengayaan Representasi Fitur.....	51
4.	Perbandingan dengan Model Alternatif.....	52
5.	Implementasi Sistem Otomatis dan Pemantauan Berkelanjutan	52
REFERENSI		53

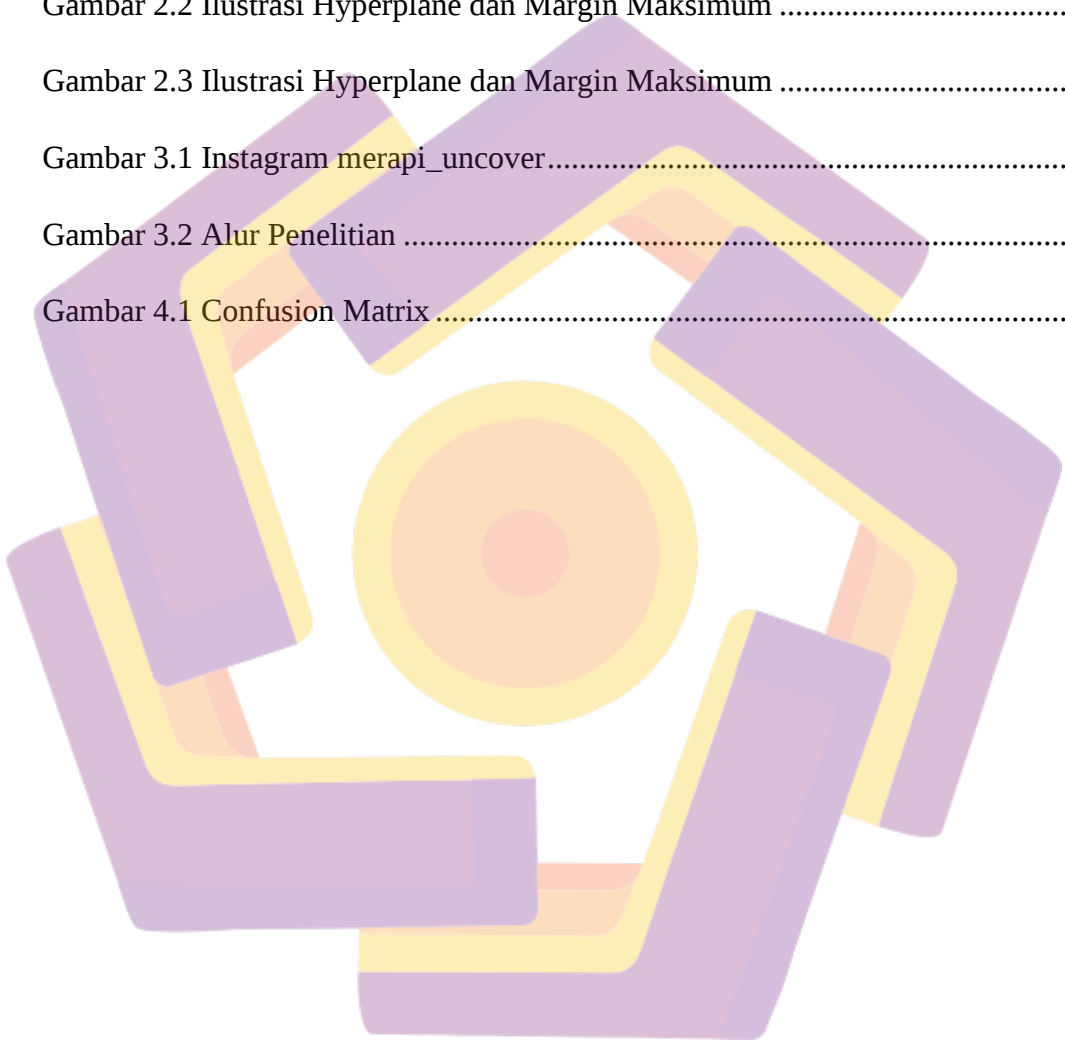


DAFTAR TABEL

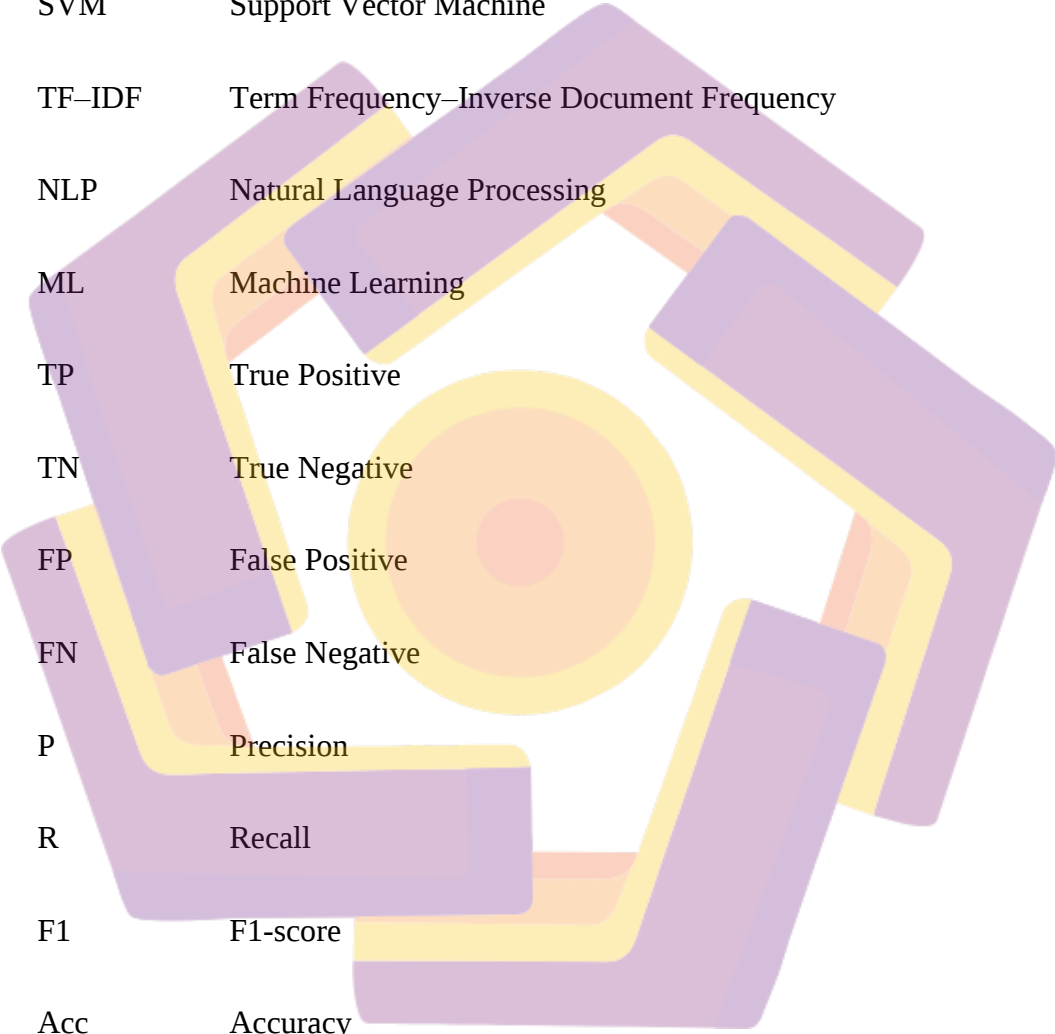
Tabel 2.1 Keaslian Penelitian	9
Tabel 3.1 Representasi TF-IDF (Sederhana).....	33
Tabel 3.2 Perbandingan dengan Kelas Lain	34
Tabel 4.1 Contoh Dataset.....	38
Tabel 4.2 Distribusi Label jenis_berita_clean.....	39
Tabel 4.3 Contoh Hasil Preprocessing Caption	41
Tabel 4.4 akumulasi evaluasi model per kategori topik.....	45

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Hyperplane dan Margin Maksimum	24
Gambar 2.2 Ilustrasi Hyperplane dan Margin Maksimum	24
Gambar 2.3 Ilustrasi Hyperplane dan Margin Maksimum	25
Gambar 3.1 Instagram merapi_uncover	28
Gambar 3.2 Alur Penelitian	29
Gambar 4.1 Confusion Matrix	44

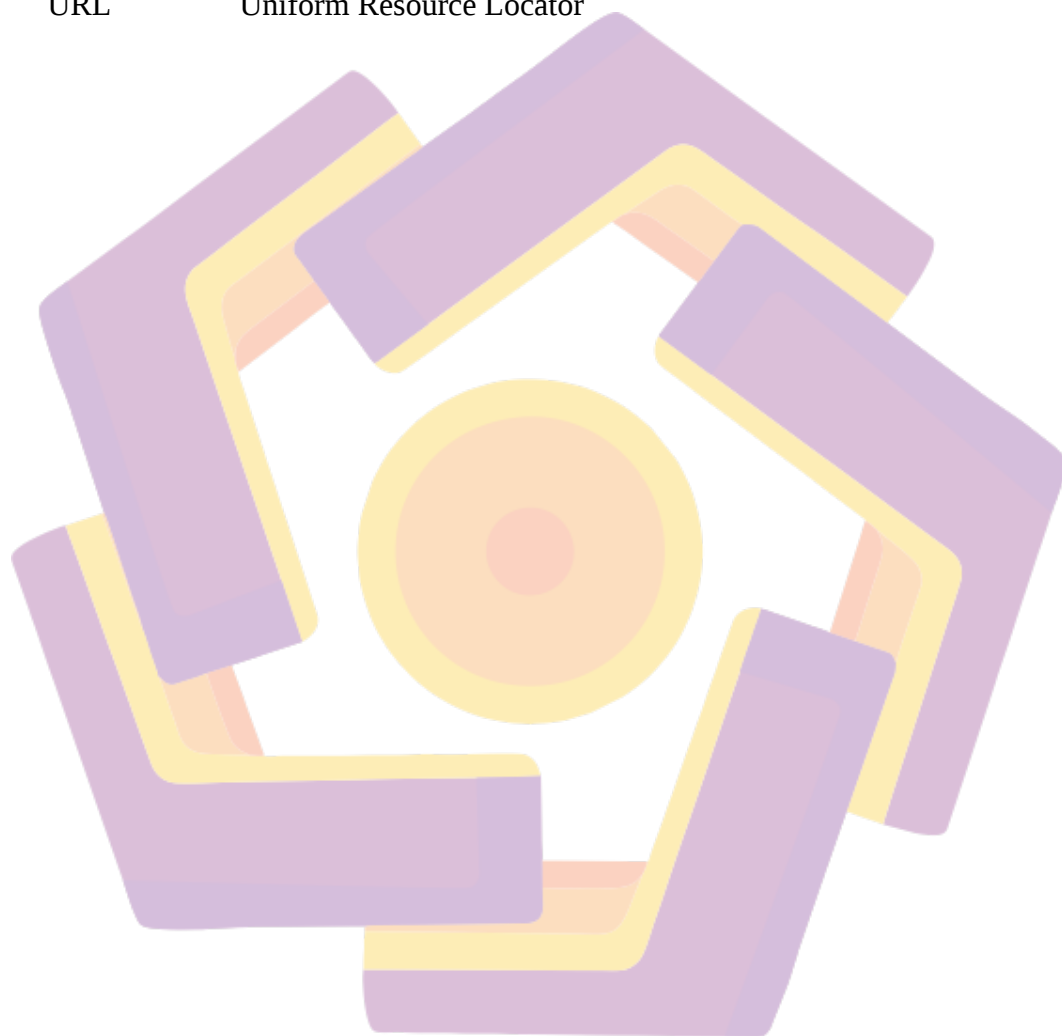


DAFTAR LAMBANG DAN SINGKATAN



SVM	Support Vector Machine
TF-IDF	Term Frequency-Inverse Document Frequency
NLP	Natural Language Processing
ML	Machine Learning
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
P	Precision
R	Recall
F1	F1-score
Acc	Accuracy
AP	Application Programming Interface
BoW	Bag of Words
SMOTE	Synthetic Minority Over-sampling Technique
LSTM	Long Short-Term Memory

BERT	Bidirectional Encoder Representations from Transformers
CNN	Convolutional Neural Network
MMH	Maximum Marginal Hyperplane
URL	Uniform Resource Locator



DAFTAR ISTILAH

Analisis sentimen

Proses menganalisis teks untuk mengidentifikasi sikap atau opini (positif, negatif, netral) terhadap suatu objek atau topik.

Caption

Teks keterangan yang menyertai unggahan (foto atau video) pada media sosial Instagram.

Cleaning teks

Tahap pembersihan teks dari elemen yang tidak relevan seperti URL, mention, hashtag, angka, tanda baca, emoji, serta normalisasi huruf dan spasi.

Feature (fitur)

Representasi numerik dari karakteristik data (misalnya bobot TF-IDF setiap kata) yang digunakan sebagai masukan model pembelajaran mesin.

Klasifikasi teks

Tugas mengelompokkan dokumen teks ke dalam satu atau beberapa kategori yang telah ditentukan.

Merapi Uncover

Akun Instagram yang berfungsi sebagai platform informasi dan berbagi cerita mengenai aktivitas Merapi, lalu lintas, kriminalitas, dan fenomena sosial di Yogyakarta.

Preprocessing teks

Serangkaian langkah awal pengolahan teks (cleaning, *stopword removal*, normalisasi) untuk menyiapkan data sebelum pemodelan.

Stopword

Kata umum yang dianggap kurang informatif (misalnya “yang”, “dan”, “di”) dan biasanya dihapus pada tahap preprocessing.

Support Vector Machine (SVM)

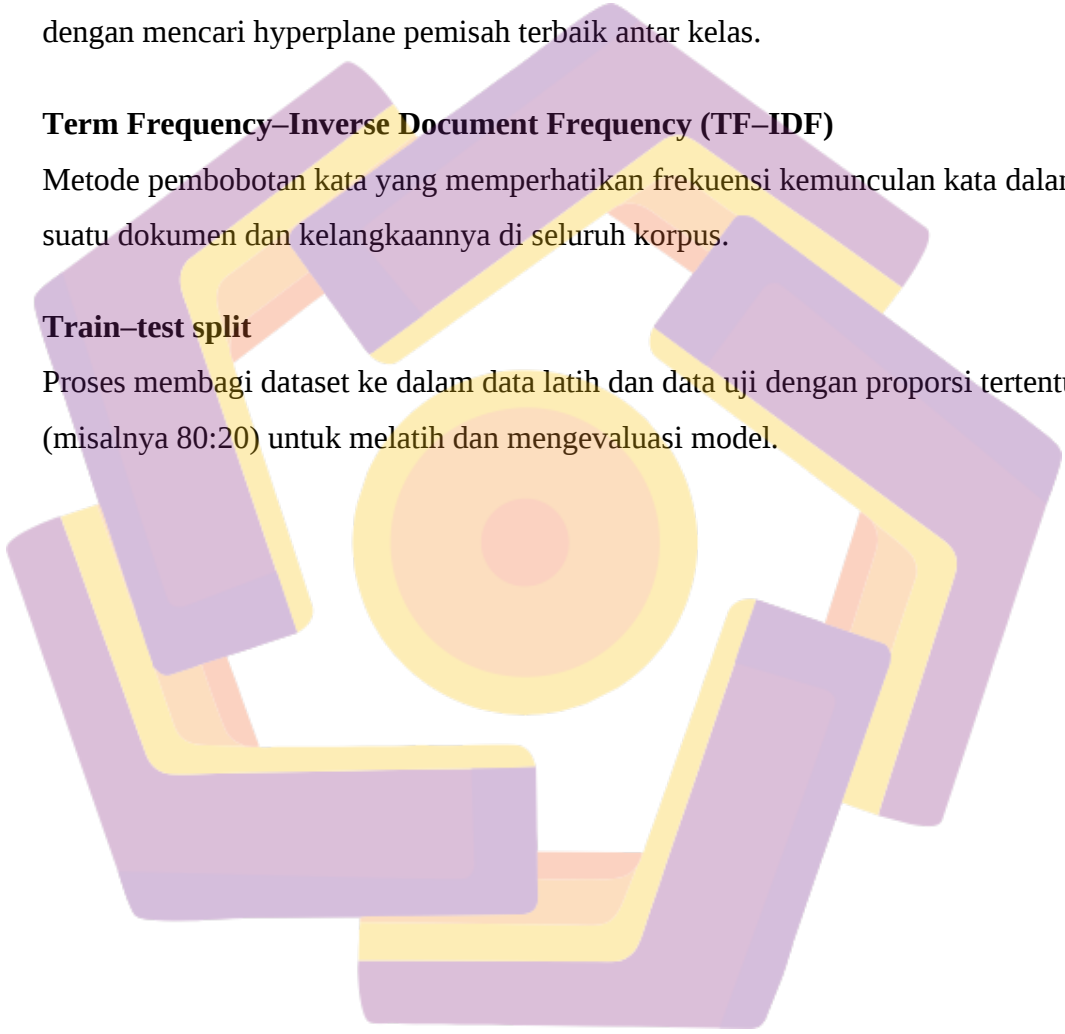
Algoritma pembelajaran mesin terawasi untuk klasifikasi dan regresi yang bekerja dengan mencari hyperplane pemisah terbaik antar kelas.

Term Frequency–Inverse Document Frequency (TF–IDF)

Metode pembobotan kata yang memperhatikan frekuensi kemunculan kata dalam suatu dokumen dan kelangkaannya di seluruh korpus.

Train–test split

Proses membagi dataset ke dalam data latih dan data uji dengan proporsi tertentu (misalnya 80:20) untuk melatih dan mengevaluasi model.



INTISARI

Perkembangan media sosial menjadikan Instagram sebagai salah satu sumber informasi utama masyarakat, termasuk dalam konsumsi berita lokal. Salah satu akun yang banyak diakses adalah Merapi Uncover yang memuat beragam informasi seputar aktivitas Gunung Merapi, lalu lintas, kriminalitas, hingga hiburan di wilayah Yogyakarta. Keragaman topik menyebabkan pengguna kesulitan menemukan berita sesuai kebutuhan, sehingga dibutuhkan sistem yang mampu melakukan klasifikasi topik secara otomatis. Penelitian ini bertujuan menganalisis performa algoritma *Support Vector Machine* (SVM) untuk klasifikasi topik berita pada akun Instagram Merapi Uncover. Dataset terdiri dari 1.501 postingan akun merapi_uncover pada periode 10–15 November 2025. Tahapan penelitian meliputi pengumpulan data melalui *web scraping*, pemetaan 77 kategori awal menjadi lima kelas besar (Hiburan / Budaya / Wisata, Lalu Lintas & Transportasi, Bencana & Kebencanaan, Sosial & Pemerintahan, serta Kriminal & Keamanan Publik), *preprocessing text* (cleaning dan *stopword removal* dengan Sastrawi), pembentukan fitur menggunakan TF-IDF, pembagian data 80:20, pelatihan model SVM linear, serta evaluasi menggunakan akurasi dan F1-score. Hasil pengujian menunjukkan bahwa model menghasilkan akurasi 81% dengan weighted F1-score 0,79. Kelas Hiburan / Budaya / Wisata memiliki F1-score tertinggi 0,91, diikuti Lalu Lintas & Transportasi sebesar 0,82, sedangkan Bencana & Kebencanaan dan Sosial & Pemerintahan masih relatif rendah. Temuan ini menunjukkan bahwa SVM dengan TF-IDF efektif digunakan untuk klasifikasi topik berita pada Merapi Uncover, meskipun peningkatan kualitas data dan penyeimbangan kelas tetap diperlukan untuk meningkatkan performa pada seluruh kategori.

Kata kunci: klasifikasi berita, Instagram, Merapi Uncover, *Support Vector Machine*, TF-IDF.

ABSTRACT

The development of social media has made Instagram one of the main sources of information for the public, including for consuming local news. One of the most accessed accounts is Merapi Uncover, which contains a variety of information about Mount Merapi's activities, traffic, crime, and entertainment in the Yogyakarta area. The diversity of topics makes it difficult for users to find news according to their needs, so a system capable of automatically classifying topics is needed. This study aims to analyze the performance of the Support Vector Machine (SVM) algorithm for classifying news topics on the Merapi Uncover Instagram account. The dataset consists of 1,501 posts from the merapi_uncover account during the period of November 10–15, 2025. The research stages include data collection through web scraping, mapping 77 initial categories into five major classes (Entertainment/Culture/Tourism, Traffic & Transportation, Disasters & Catastrophes, Social & Government, and Crime & Public Safety), text preprocessing (cleaning and stopword removal with Sastrawi), feature formation using TF–IDF, 80:20 data division, linear SVM model training, and evaluation using accuracy and F1-score. The test results show that the model produces an accuracy of 81% with a weighted F1-score of 0.79. The Entertainment/Culture/Tourism class has the highest F1-score of 0.91, followed by Traffic & Transportation at 0.82, while Disaster & Disaster Management and Social & Government are still relatively low. These findings indicate that SVM with TF–IDF is effective for classifying news topics on Merapi Uncover, although improvements in data quality and class balancing are still needed to enhance performance across all categories.

Keywords: news classification, Instagram, Merapi Uncover, Support Vector Machine, TF–IDF, Five