

**KOMPARASI KINERJA ALGORITMA NAIVE BAYES DAN
SUPPORT VECTOR MACHINE (SVM) DALAM
KLASIFIKASI SENTIMEN WARGANET TWITTER
MENGENAI ISU IBU KOTA NUSANTARA (IKN) PADA
SEKTOR PENDIDIKAN**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

EGI APREZA

22.12.2483

Kepada

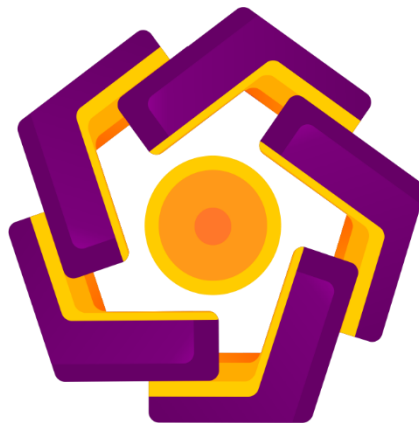
**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**KOMPARASI KINERJA ALGORITMA NAIVE BAYES DAN
SUPPORT VECTOR MACHINE (SVM) DALAM
KLASIFIKASI SENTIMEN WARGANET TWITTER
MENGENAI ISU IBU KOTA NUSANTARA (IKN) PADA
SEKTOR PENDIDIKAN**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

EGI APREZA

22.12.2483

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

yang disusun dan diajukan oleh

EGI APREZA

22.12.2483

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Februari 2026

Dosen Pembimbing,



Anggit Dwi Hartanto, S.Kom., M.Kom

NIK. 190302163

HALAMAN PENGESAHAN
SKRIPSI
KOMPARASI KINERJA ALGORITMA NAIVE BAYES DAN SUPPORT
VECTOR MACHINE (SVM) DALAM KLASIFIKASI SENTIMEN
WARGANET TWITTER MENGENAI ISU IBU KOTA NUSANTARA (IKN)
PADA SEKTOR PENDIDIKAN

yang disusun dan diajukan oleh
EGI APREZA

22.12.2483

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Sharazita Dyah Anggita, S.Kom., M.Kom.
NIK. 190302285

Ninik Tri Hartanti, S.Kom., M.Kom.
NIK. 190302330

Anggit Dwi Hartanto, S.Kom., M.Kom.
NIK. 190302163



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
24 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : EGI APREZA
NIM : 22.12.2483

Menyatakan bahwa Skripsi dengan judul berikut:

KOMPARASI KINERJA ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI SENTIMEN WARGANET TWITTER MENGENAI ISU IBU KOTA NUSANTARA (IKN) PADA SEKTOR PENDIDIKAN

Dosen Pembimbing : Anggit Dwi Hartanto, S.Kom., 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, 24 Februari 2026

Yang Menyatakan



EGI APREZA

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan dengan penuh rasa syukur kepada Allah SWT atas segala rahmat, kesehatan, kekuatan, serta kemudahan yang diberikan sehingga saya dapat menyelesaikan skripsi ini dengan baik. Proses penyusunan skripsi ini bukanlah perjalanan yang mudah, melainkan perjalanan yang dipenuhi berbagai tantangan, keterbatasan, serta suka dan duka yang menjadi bagian dari proses pembelajaran, kesabaran, dan pendewasaan diri.

Dengan segala kerendahan hati dan rasa cinta yang tak terhingga, skripsi ini saya persembahkan kepada kedua orang tua tercinta. Terima kasih atas setiap doa yang tidak pernah putus, setiap pengorbanan yang tidak pernah diperhitungkan, serta kasih sayang yang selalu menguatkan langkah saya hingga berada pada titik ini. Segala dukungan, kepercayaan, nasihat, dan ketulusan yang diberikan menjadi sumber kekuatan terbesar dalam perjalanan hidup dan pendidikan saya. Tiada kata yang mampu menggambarkan betapa besar rasa cinta, hormat, dan terima kasih saya kepada kalian.

Saya juga menyampaikan terima kasih kepada dosen pembimbing yang telah dengan sabar memberikan bimbingan, arahan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini. Ucapan terima kasih juga saya sampaikan kepada seluruh dosen dan civitas akademika atas ilmu, pengalaman, dan pembelajaran berharga selama masa perkuliahan.

Skripsi ini juga saya persembahkan kepada keluarga, sahabat, serta semua pihak yang telah memberikan doa, dukungan, dan semangat dalam perjalanan akademik saya hingga karya ini dapat terselesaikan.

Dan secara khusus, skripsi ini saya persembahkan kepada seseorang yang tidak dapat saya sebutkan namanya. Seseorang yang dengan tulus telah menjadi bagian dari perjalanan dan semangat saya hingga sampai pada titik ini. Terima kasih atas dukungan, perhatian, kesabaran, serta ketulusan yang selalu diberikan dalam setiap proses yang saya lalui.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta
2. Bapak Hanif Al-Fatta, M.Kom, Ph.D selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto, M.Kom., selaku Ketua Program Studi Sistem Informasi sekaligus dosen pembimbing yang telah memberikan arahan, saran, dan motivasi kepada penulis selama proses penyusunan tugas akhir ini.
4. Kedua orang tua, keluarga besar, dan teman-teman tercinta yang memberikan semangat dan doa kepada penulis.
5. Seluruh pihak lain yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak langsung selama proses penyusunan laporan tugas akhir ini.

Yogyakarta, 24 Febuari 2026



Egi Apreza

DAFTAR ISI

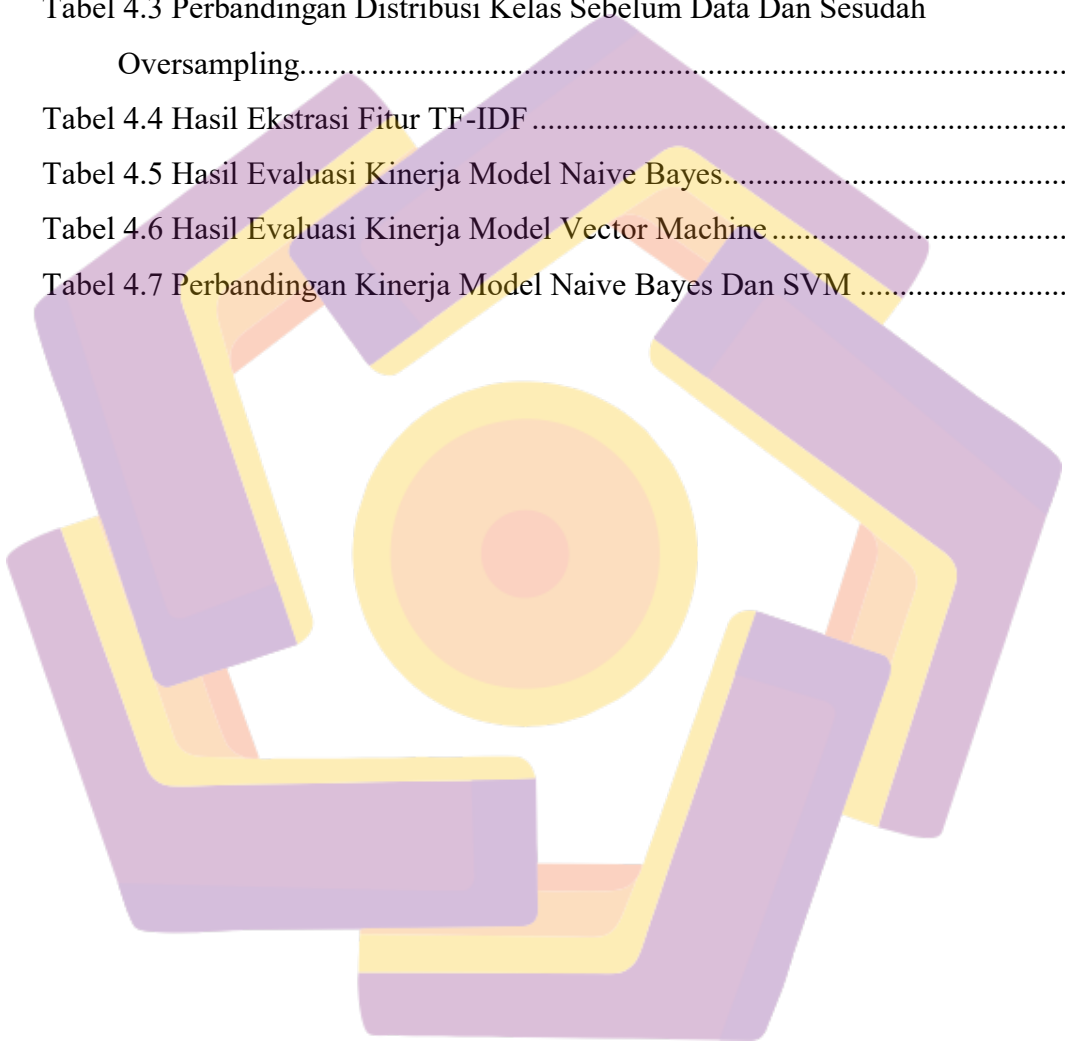
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH	xv
INTISARI.....	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Studi Literatur.....	8
2.2 Dasar Teori	12
2.2.1 Ibu Kota Nusantara (IKN) dan Sektor Pendidikan	12
2.2.2 Analisis Sentimen (<i>Sentiment Analysis</i>)	12
2.2.3 Klasifikasi Sentimen (<i>Sentiment Classification</i>)	12
2.2.4 Tahapan Pra-pemrosesan Teks (<i>Text Preprocessing</i>).....	13
2.2.5 Ekstraksi Fitur: TF-IDF	13

2.2.6	Twitter sebagai Sumber Data	14
2.2.7	Algoritma Klasifikasi : Algoritma Naive Bayes (NB)	14
2.2.8	Algoritma Klasifikasi : Algoritma Support Vector Machine	14
2.3	Bahasa Pemrograman dan Library	15
2.3.1	Python	15
2.3.2	Library Pendukung	15
2.4	Metode Pengujian dan Evaluasi	16
2.4.1	Confusion Matrix	16
2.4.2	Metrik Evaluasi Kinerja	16
2.4.3	Cross Validation	17
BAB III METODE PENELITIAN		18
3.1	Objek Penelitian	18
3.2	Alur Penelitian	18
3.2.1	Pengumpulan Data	20
3.2.2	Pra-pemrosesan Data (<i>Text Preprocessing</i>)	20
3.2.3	Pelabelan Sentimen Otomatis	20
3.2.4	Penyeimbangan Data (<i>Oversampling</i>)	20
3.2.5	Pembagian Data	21
3.2.6	Ekstraksi Fitur (<i>TF-IDF</i>)	21
3.2.7	Pemodelan (<i>Modeling</i>)	22
3.2.8	Evaluasi dan Validasi (<i>Evaluation & Validation</i>)	22
3.2.9	Perbandingan Kinerja Model	24
3.3	Alat dan Bahan	25
3.3.1	Data Penelitian	25
3.3.2	Alat dan Instrumen Penelitian	25
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Hasil Penelitian	27
4.1.1	Pengumpulan Data Twitter	27
4.1.2	Hasil Pra-pemrosesan Data	28
4.1.3	Tokenisasi Per Kalimat	31

4.1.4 Pelabelan Sentimen Otomatis IndoBERT	32
4.1.5 Penyeimbang Data (Oversampling)	33
4.1.6 Pembagian Data Training Dan Testing	38
4.1.7 Ekstraksi Fitur Menggunakan TF-IDF	39
4.1.8 Training Model Naive Bayes Dan SVM	40
4.1.9 Evaluasi Akurasi Model	41
4.1.10 Perbandingan Kerja Model.....	46
4.1.11 Pembahasan Hasil Penelitian	48
4.2 Implementasi Sistem Dan Aplikasi Streamlit	49
4.2.1 Implementasi Script Pada Visual Studio Code	49
4.2.2 Hasil Implementasi Aplikasi Streamlit	52
4.2.3 Pembahasan Hasil Implementasi.....	56
BAB V PENUTUP.....	57
5.1 Kesimpulan	57
5.2 Saran	58
REFERENSI	60

DAFTAR TABEL

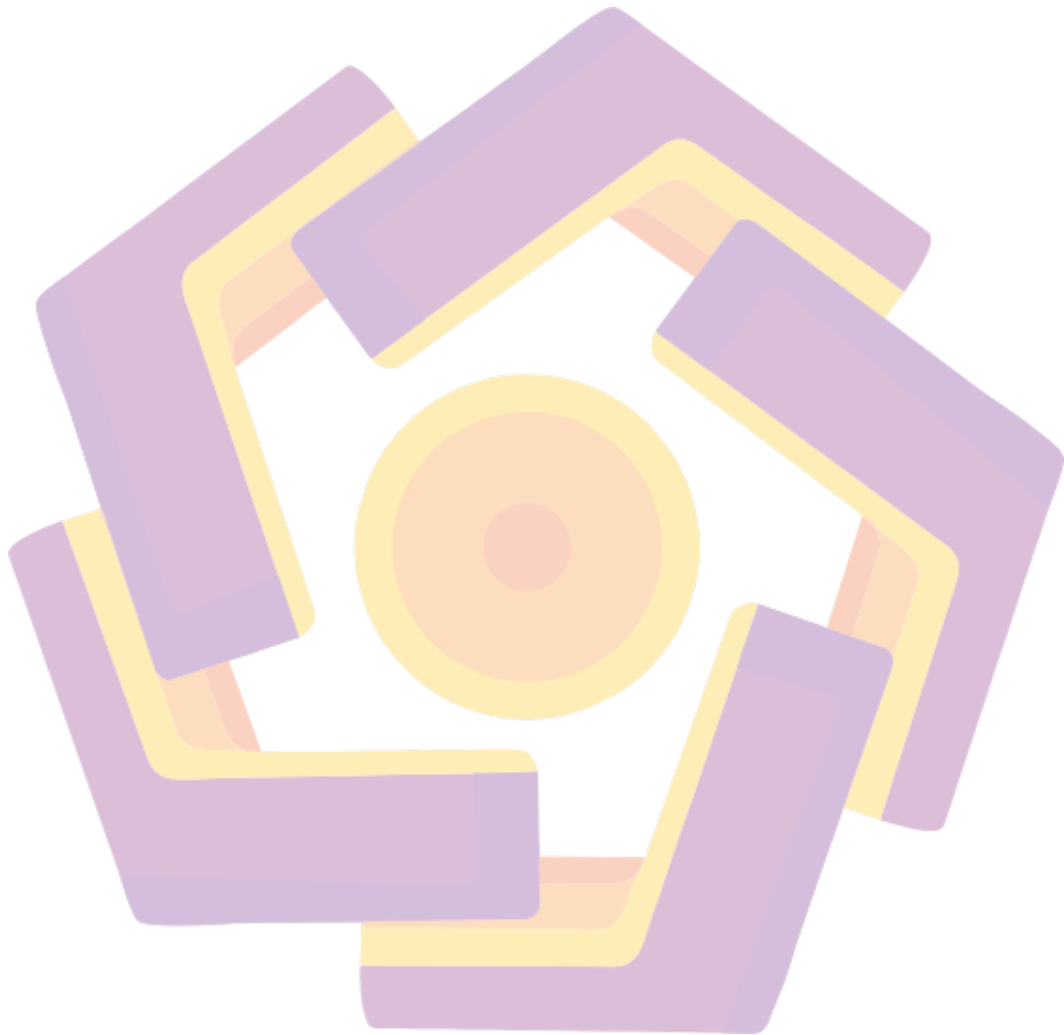
Tabel 2.1 Keaslian Penelitian.....	9
Tabel 4.1 Jumlah Data Tweet Hasil Pengumpulan	27
Tabel 4.2 Distribusi Data Sentiment Setelah Pelabelan.....	32
Tabel 4.3 Perbandingan Distribusi Kelas Sebelum Data Dan Sesudah Oversampling.....	34
Tabel 4.4 Hasil Ekstraksi Fitur TF-IDF	40
Tabel 4.5 Hasil Evaluasi Kinerja Model Naive Bayes.....	42
Tabel 4.6 Hasil Evaluasi Kinerja Model Vector Machine	43
Tabel 4.7 Perbandingan Kinerja Model Naive Bayes Dan SVM	46



DAFTAR GAMBAR

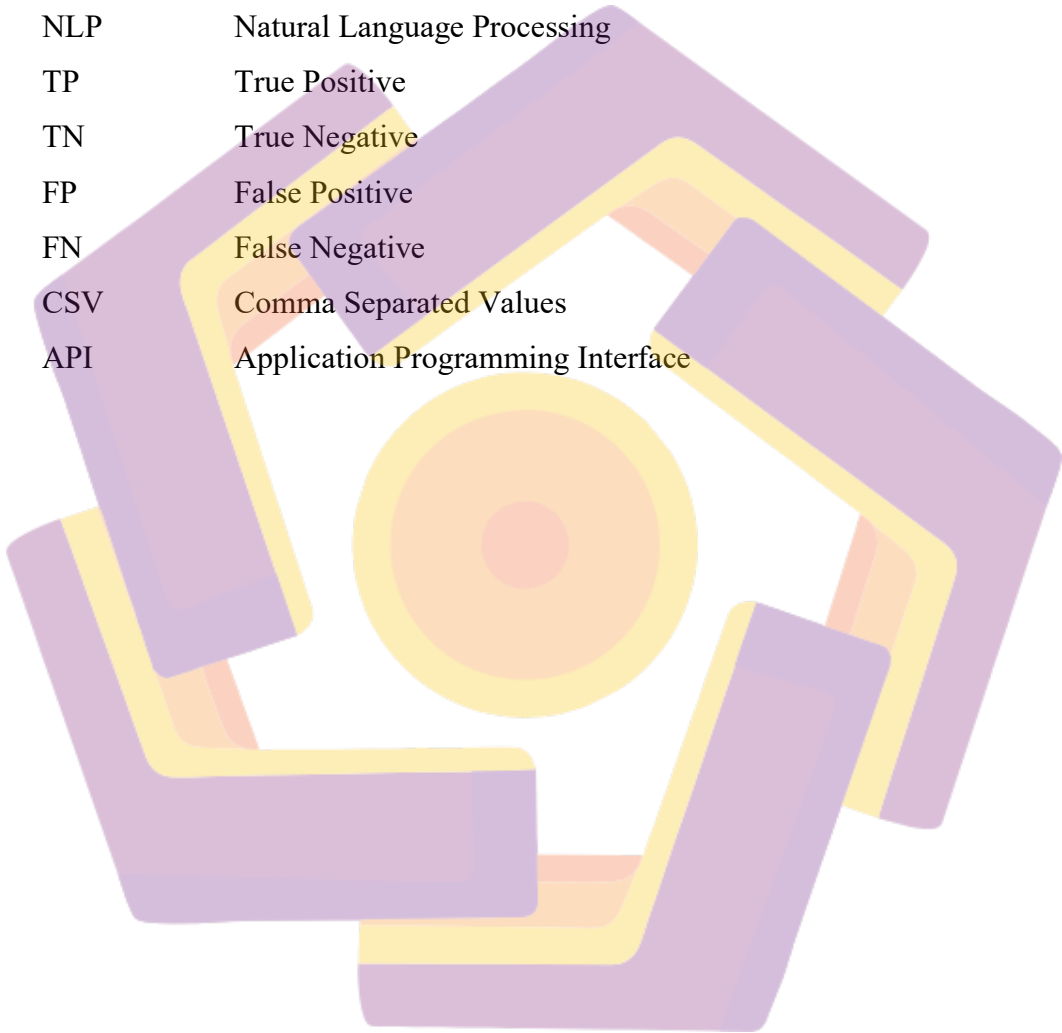
Gambar 3.1 Alur Penelitian	19
Gambar 4.1 Hasil Sebelum Dan Sesudah Pra-Pemrosesan Data	28
Gambar 4.2 Grafik Top 10 Most Frequent Tokens.....	29
Gambar 4.3 Word Cloud Token Hasil Preprocessing.....	30
Gambar 4.4 Hasil Tokenisasi Per Kalimat.....	31
Gambar 4.5 Visualisasi Distribusi Sentimen	35
Gambar 4.6 Word Cloud Sentimen Positif	36
Gambar 4.7 Word Cloud Sentimen Negatif.....	37
Gambar 4.8 Word Cloud Sentimen Netral.....	37
Gambar 4.9 Pembagian Data Training Dan Testing.....	39
Gambar 4.10 Algoritma Naive Bayes	40
Gambar 4.11 Algoritma Support Vector Machine (SVM)	41
Gambar 4.12 Confusion Matrix Model Naive Bayes	44
Gambar 4.13 Confusion Matrix Model SVM	45
Gambar 4.14 Grafik Perbandingan Akurasi Model	47
Gambar 4.15 Script Pemanggilan Model Klasifikasi Sentimen	50
Gambar 4.16 Script Tahapan Pre Processing.....	50
Gambar 4.17 Script Proses Prediksi Sentimen.....	52
Gambar 4.18 Tampilan Awal Aplikasi Streamlit.....	53
Gambar 4.19 Hasil Analisis Kalimat Positif.....	54
Gambar 4.20 Hasil Analisis Kalimat Negatif	54
Gambar 4.21 Hasil Analisis Kalimat Netral	55

DAFTAR LAMPIRAN



DAFTAR LAMBANG DAN SINGKATAN

IKN	Ibu Kota Nusantara
NB	Naive Bayes
SVM	Support Vector Machine
TF-IDF	Term Frequency – Inverse Document Frequency
NLP	Natural Language Processing
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
CSV	Comma Separated Values
API	Application Programming Interface



DAFTAR ISTILAH

Analisis Sentimen	Klasifikasi opini teks menjadi positif, negatif, atau netral.
Crawling	Pengambilan data otomatis dari media sosial.
Preprocessing	Proses pembersihan dan normalisasi teks.
Tokenisasi	Pemecahan teks menjadi kata-kata.
Stemming	Pengubahan kata menjadi bentuk dasar.
Confusion Matrix	Tabel evaluasi hasil klasifikasi.
Cross-Validation	Teknik pengujian model mengukur kestabilan kinerja.
Oversampling	Teknik penyeimbangan jumlah data antar kelas.

INTISARI

Pembangunan Ibu Kota Nusantara (IKN) sebagai proyek strategis nasional menimbulkan beragam respons dari masyarakat, khususnya terkait dampaknya terhadap sektor pendidikan. Media sosial Twitter (X) menjadi sarana utama bagi warganet untuk menyampaikan opini, baik yang bersifat mendukung, menolak, maupun netral. Banyaknya data opini yang muncul secara cepat dan masif menyebabkan analisis manual menjadi tidak efektif, sehingga diperlukan pendekatan komputasi untuk mengidentifikasi dan mengukur kecenderungan sentimen publik secara objektif. Ketidakstabilan dan tingginya variasi sentimen tersebut berpotensi memengaruhi persepsi publik serta pengambilan kebijakan di sektor pendidikan.

Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma Naive Bayes dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen warganet Twitter terkait isu IKN pada sektor pendidikan. Data penelitian berupa 2.006 tweet berbahasa Indonesia yang dikumpulkan melalui proses crawling. Tahapan penelitian meliputi pra-pemrosesan teks (cleaning, case folding, tokenisasi, stopword removal, dan stemming), pelabelan sentimen otomatis menggunakan IndoBERT, penyeimbangan data dengan oversampling, ekstraksi fitur menggunakan TF-IDF, serta evaluasi model menggunakan confusion matrix dan Stratified K-Fold Cross-Validation.

Kata kunci: Analisis Sentimen, Twitter, Ibu Kota Nusantara, Pendidikan, Naive Bayes, Support Vector Machine

ABSTRACT

The development of the Nusantara Capital City (IKN) as a national strategic project has generated diverse public responses, particularly regarding its impact on the education sector. Social media platform Twitter (X) has become a primary medium for netizens to express their opinions, whether supportive, critical, or neutral. The rapid and massive emergence of opinion data renders manual analysis ineffective, thereby necessitating computational approaches to objectively identify and measure public sentiment trends. The instability and high variability of these sentiments have the potential to influence public perception as well as policymaking in the education sector.

This study aims to analyze and compare the performance of the Naive Bayes and Support Vector Machine (SVM) algorithms in classifying Twitter users' sentiments related to the IKN issue in the education sector. The research data consist of 2,006 Indonesian-language tweets collected through a crawling process. The research stages include text preprocessing (cleaning, case folding, tokenization, stopword removal, and stemming), automated sentiment labeling using IndoBERT, data balancing through oversampling, feature extraction using TF-IDF, and model evaluation using a confusion matrix and Stratified K-Fold Cross-Validation.

Keyword: Sentiment Analysis, Twitter, Nusantara Capital City (IKN), Education, Naive Bayes, Support Vector Machine