

**IMPLEMENTASI METODE SUPPORT VECTOR MACHINE
DAN NAIVE BAYES UNTUK ANALISIS OPINI PUBLIK
TENTANG HANTAVIRUS PADA PLATFORM TIKTOK**

TUGAS AKHIR

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer



MIFTAHUL JANNAH

1222001001

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE**

2026

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Miftahul Jannah

NIM : 1222001001

Tanda Tangan :



Tanggal : 21 September 2026

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Miftahul Jannah
NIM : 1222001001
Program Studi : Teknik Informatika
Fakultas : Teknik dan Ilmu Komputer
Judul Penelitian : Implementasi Metode Support Vector Machine dan Naive Bayes untuk Analisis Opini Publik Tentang Hantavirus pada Platform TikTok

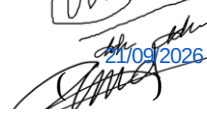
Telah berhasil menyelesaikan tugas akhir dan diterima sebagai bagian persyaratan yang diperlukan dalam mengikuti sidang tugas akhir untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

DEWAN PENGUJI

Pembimbing 1 : Ir. Iwan Adhicandra, S.T., M.Sc., Ph.D., SMIEEE

Pembahas 1 : Dewi Fatmawati Surianto, S.Kom., M.Kom.

Pembahas 2 : Dimas Aryo Anggoro, S.Kom., M.Sc.


Ditetapkan di Jakarta

Tanggal 18 September 2026

KATA PENGANTAR

Puji syukur yang dipanjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat -Nya, penulis dapat menyelesaikan Tugas Akhir ini dengan judul IMPLEMENTASI METODE SUPPORT VECTOR MACHINE DAN NAIVE BAYES UNTUK ANALISIS OPINI PUBLIK TENTANG HANTAVIRUS PADA PLATFORM TIKTOK. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Program Studi Teknik Informatika pada Fakultas Teknik dan Ilmu Komputer Universitas Bakrie. Diantara seluruh halaman dalam skripsi ini, lembar ungkapan terimakasih adalah halaman yang paling bermakna, karena disinilah penulis menitipkan rasa terimakasih yang tidak mampu terucap oleh kata-kata. Bismillahirrahmanirrahim, skripsi ini penulis persembahkan kepada:

- 1) Ibu Prof. Ir. Sofia W. Alisjahbana, M.Sc., Ph.D., IPU., ASEAN Eng., selaku Rektor Universitas Bakrie.
- 2) Bapak Dr. Mohammad Ihsan, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Ilmu Komputer Universitas Bakrie.
- 3) Kepada Bapak Ir. Iwan Adhichandra, S.T., M.Sc., Ph.D., SMIEEE selaku Dosen Pembimbing. Terimakasih yang sebesar-besarnya atas waktu, kesabaran dan dukungan yang Bapak berikan selama proses penulisan skripsi ini. Tanpa bimbingan dan dukungan bapak mungkin skripsi ini belum bisa selesai. Terimakasih telah mempermudah setiap proses, selalu membuka pintu konsultasi dengan lapang. Bapak bukan hanya membimbing dengan ilmu, tetapi juga dengan hati. Semoga Allah SWT senantiasa melimpahkan kesehatan, keberkahan, serta balasan yang terbaik untuk setiap kebaikan yang Bapak berikan.
- 4) Kedua orangtua tercinta, Bapak Emriyal dan Ibu Roslaini, dua sosok hebat yang menjadi alasan utama penulis berjuang hingga titik ini. Meskipun kita bukan keluarga yang terbiasa berbicara dari hati ke hati, penulis tahu dan merasakan doa dan harapan yang tak henti mengalir dibalik setiap keheningan. Penulis sadar bahwa cinta kalian mewujud dalam bentuk yang berbeda, namun nyata dalam tindakan dan kehadiran yang menguatkan. Terimakasih telah menjadi sandaran dan sumber kebahagiaan bagi penulis. Skripsi ini hanyalah awal dari upaya untuk membuat kalian bangga.
- 5) Saudara laki-laki penulis, Abang Heru Vernando. Terimakasih atas pengorbanan dan segala usaha yang membantu penulis dalam mendapatkan gelar sarjana komputer.

Terimakasih telah mengorbankan mimpi-mimpimu untuk membiayai pendidikanku. Setiap tetes keringatmu dalam mencari nafkah adalah bukti kasih sayang yang tak ternilai. Keberhasilanku hari ini adalah berkat dari pengorbananmu.

- 6) Seseorang yang tidak kalah penting, kaka Siska. Terimakasih telah menjadi bagian dalam proses perjalanan penulis menyusun skripsi. Berkontribusi baik tenaga, waktu, menemani, mendukung, serta menghibur penulis dalam kesedihan, mendengarkan keluh kesah dan meyakinkan penulis untuk pantang menyerah hingga penyusunan skripsi ini terselesaikan.
- 7) Teman-teman seperjuangan, khususnya Andy, Ilham, Killy, Nofita, Pajar, dan Theresia yang selalu kebersamai dalam proses skripsi ini, saling menguatkan dan berbagi tawa. Terimakasih untuk cerita dan semangat yang tidak pernah padam.
- 8) *Last but not least*, anak perempuan terakhir dan harapan terakhir orangtuanya. Ya, penulis sendiri. Apresiasi untuk diriku sendiri yang mungkin tidak pandai menunjukkan perasaan, tetapi tetap memilih untuk bertahan, berjalan, dan menyelesaikan apa yang sudah dimulai. Terimakasih sudah kuat dan sudah tidak menyerah, meskipun tidak selalu yakin. Semoga langkah ini menjadi bukti, bahwa penulis mampu.

Jakarta, April 2026



Miftahul Jannah

NIM 1222001001

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

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

Nama : Miftahul Jannah

NIM : 1222001001

Program Studi : Teknik Informatika

Fakultas : Teknik dan Ilmu Komputer

Jenis Tugas Akhir : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul:

IMPLEMENTASI METODE SUPPORT VECTOR MACHINE DAN NAIVE BAYES UNTUK ANALISIS OPINI PUBLIK TENTANG HANTAVIRUS PADA PLATFORM TIKTOK.

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Noneksklusif ini Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, 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 untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 21 September 2026

Yang
menyatakan



Miftahul Jannah

DAFTAR ISI

<u>HALAMAN HALAMAN PERNYATAAN ORISINALITAS</u>	i
<u>HALAMAN PENGESAHAN</u>	ii
<u>KATA PENGANTAR</u>	iii
<u>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI</u>	v
<u>DAFTAR ISI</u>	vi
<u>ABSTRAK</u>	ix
<u>ABSTRACT</u>	x
<u>DAFTAR GAMBAR</u>	xi
<u>DAFTAR TABEL</u>	xii
<u>DAFTAR LAMPIRAN</u>	xiii
<u>BAB I</u>	1
<u>PENDAHULUAN</u>	1
<u>1.1 Latar Belakang</u>	1
<u>1.2 Rumusan Masalah</u>	8
<u>1.3 Tujuan Penelitian</u>	8
<u>1.4 Manfaat Penelitian</u>	9
<u>1.4.1 Manfaat Teoritis</u>	9
<u>1.4.2 Manfaat Praktis</u>	9
<u>1.5 Batasan Masalah</u>	10
<u>1.6 Sistematika Penulisan</u>	11
<u>BAB II</u>	13
<u>TINJAUAN PUSTAKA</u>	13
<u>2.1 Kajian Penelitian</u>	13
<u>2.2 Landasan Teori</u>	21
<u>2.2.1 Analisis Sentimen</u>	21
<u>2.2.2 TikTok sebagai Platform Media Sosial</u>	22
<u>2.2.3 Natural Language Processing (NLP)</u>	23
<u>2.2.4 Text Preprocessing</u>	24
<u>2.2.5 Support Vector Machine (SVM)</u>	25
<u>2.2.6 Term Frequency–Inverse Document Frequency (TF-IDF)</u>	27
<u>2.2.7 Naive Bayes</u>	29
<u>2.2.8 Confusion Matrix</u>	30
<u>2.2.9 Evaluasi Model</u>	32
<u>2.3 Kerangka Berpikir</u>	35
<u>BAB III</u>	37

<u>METODOLOGI PENELITIAN</u>	37
<u>3.1 Tahapan Penelitian</u>	37
<u>3.2. Metode Penelitian</u>	38
<u>3.3 Pengumpulan Data</u>	39
<u>3.3.1 Sumber Data</u>	39
<u>3.3.2 Proses Pengumpulan Data</u>	42
<u>3.4 Pengolahan Data</u>	43
<u>3.4.1 Filtering Bahasa Daerah</u>	44
<u>3.4.2 Text Preprocessing</u>	44
<u>3.4.3 Pelabelan Sentimen</u>	48
<u>3.4.4 Ekstraksi Fitur (TF-IDF)</u>	50
<u>3.5 Pembangunan Model</u>	50
<u>3.5.1 Train-Test Split</u>	51
<u>3.5.2 Linear Support Vector Classifier (LinearSVC)</u>	52
<u>3.5.3 Multinomial Naive Bayes</u>	52
<u>3.6 Evaluasi Model</u>	53
<u>3.7 Perangkat Penelitian</u>	54
<u>3.7.1 Perangkat Keras</u>	54
<u>3.7.2 Perangkat Lunak</u>	54
<u>BAB IV</u>	56
<u>HASIL DAN PEMBAHASAN</u>	56
<u>4.1 Hasil Pengumpulan Data</u>	56
<u>4.2 Hasil Text Preprocessing</u>	57
<u>4.2.1 Cleaning</u>	58
<u>4.2.2 Case Folding</u>	58
<u>4.2.3 Normalization</u>	59
<u>4.2.4 Tokenizing</u>	60
<u>4.2.5 Stopword Removal</u>	61
<u>4.2.6 Stemming</u>	63
<u>4.3 Pelabelan Sentimen</u>	64
<u>4.4 Hasil Analisis Sentimen Opini Masyarakat</u>	65
<u>4.5 Ekstraksi Fitur TF-IDF</u>	66
<u>4.6 Train-Test Split</u>	67
<u>4.7 Penanganan Ketidakseimbangan Data</u>	68
<u>4.8 Klasifikasi dan Evaluasi Model</u>	68
<u>4.8.1 Klasifikasi menggunakan LinearSVC</u>	69
<u>4.8.2 Klasifikasi menggunakan Multinomial Naive Bayes</u>	69

<u>4.9 Perbandingan Performa Model</u>	71
<u>4.10 Pembahasan</u>	72
<u>BAB V</u>	74
<u>KESIMPULAN DAN SARAN</u>	74
<u>5.1 Kesimpulan</u>	74
<u>5.2 Saran</u>	75
<u>DAFTAR PUSTAKA</u>	76
<u>LAMPIRAN</u>	83

“Implementasi Metode Support Vector Machine dan Naive Bayes untuk Analisis Opini Publik tentang Hantavirus pada Platform Tiktok”

Miftahul Jannah

1222001001@student.bakrie.ac.id

ABSTRAK

Hantavirus merupakan salah satu isu kesehatan yang mendapat perhatian masyarakat Indonesia pada tahun 2026 setelah ditemukan sejumlah kasus di beberapa wilayah. Tanggapan masyarakat terhadap isu tersebut banyak disampaikan melalui kolom komentar TikTok, namun volume komentar yang besar dan bersifat informal menyulitkan analisis secara manual. Penelitian ini bertujuan untuk menganalisis sentimen opini masyarakat terhadap Hantavirus pada platform TikTok serta membandingkan performa algoritma LinearSVC dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen tersebut. Data penelitian berupa 5.203 komentar dari delapan unggahan TikTok akun SCTV dan Liputan6 yang dikumpulkan melalui teknik web scraping pada periode Mei–Juni 2026. Setelah melalui tahap filtering bahasa daerah dan text preprocessing (*cleaning, case folding, normalization, tokenizing, stopword removal, dan stemming*), diperoleh 4.818 komentar yang diberi label sentimen positif, negatif, dan netral menggunakan pendekatan rule-based berbasis InSet Lexicon. Data selanjutnya direpresentasikan menggunakan TF-IDF, ditangani ketidakseimbangan kelasnya dengan Random Oversampling, dan dibagi dengan rasio 80:20 sebelum diklasifikasikan menggunakan LinearSVC dan Multinomial Naive Bayes. Hasil analisis sentimen menunjukkan bahwa opini masyarakat didominasi sentimen negatif (44,77%), diikuti sentimen positif (28,21%) dan netral (27,02%). Hasil pengujian model menunjukkan bahwa LinearSVC memperoleh performa lebih baik dibandingkan Multinomial Naive Bayes, dengan nilai accuracy 83,71%, precision 84,55%, recall 83,71%, dan F1-score 83,92%, sedangkan Multinomial Naive Bayes memperoleh accuracy 75,00%, precision 75,99%, recall 75,00%, dan F1-score 74,84%. Dengan demikian, LinearSVC merupakan algoritma dengan performa klasifikasi terbaik pada penelitian ini.

Kata kunci: Analisis Sentimen, Hantavirus, TikTok, LinearSVC, Multinomial Naive Bayes, TF-IDF

“Implementasi Metode Support Vector Machine dan Naive Bayes untuk Analisis Opini Publik tentang Hantavirus pada Platform Tiktok”

Miftahul Jannah

1222001001@student.bakrie.ac.id

ABSTRACT

Hantavirus is a health issue that attracted public attention in Indonesia in 2026 following the identification of several cases in various regions. Public responses to this issue were widely expressed through TikTok comment sections; however, the large volume of comments and their informal nature make manual analysis challenging. This study aims to analyze public sentiment regarding Hantavirus on the TikTok platform and compare the performance of the LinearSVC and Multinomial Naive Bayes algorithms in classifying these sentiments. The research data consisted of 5,203 comments collected from eight TikTok posts published by the SCTV and Liputan6 accounts through web scraping during May–June 2026. Following regional language filtering and text preprocessing (cleaning, case folding, normalization, tokenizing, stopword removal, and stemming), 4,818 comments were obtained and labeled as positive, negative, or neutral using a rule-based approach based on the InSet Lexicon. The data were subsequently represented using TF-IDF, balanced using Random Oversampling, and divided at an 80:20 ratio before being classified using LinearSVC and Multinomial Naive Bayes. The sentiment analysis results indicate that public opinion was dominated by negative sentiment (44.77%), followed by positive sentiment (28.21%) and neutral sentiment (27.02%). Model evaluation showed that LinearSVC achieved better performance than Multinomial Naive Bayes, with an accuracy of 83.71%, precision of 84.55%, recall of 83.71%, and F1-score of 83.92%. Meanwhile, Multinomial Naive Bayes achieved an accuracy of 75.00%, precision of 75.99%, recall of 75.00%, and F1-score of 74.84%. Therefore, LinearSVC demonstrated better classification performance in this study.

Keywords: Sentiment Analysis, Hantavirus, TikTok, LinearSVC, Multinomial Naive Bayes, TF-IDF.

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Berpikir.....	43
Gambar 3. 1 Flowchart Tahapan Penelitian.....	44
Gambar 4. 1 Hasil Analisis Sentimen	61
Gambar 4. 2 <i>Confusion Matrix</i> LinearSVC dan Multinomial Naive Bayes	65
Gambar 4. 3 Perbandingan Kinerja LinearSVC dan Multinomial Naive Bayes	66

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian.....	14
Tabel 2.2 Kategori Sentimen	22
Tabel 2.3 Komponen <i>Confusion Matrix</i>	31
Tabel 2.4 <i>Confusion Matrix</i> Tiga Kelas.....	32
Tabel 3.1 Sumber Data Penelitian	40
Tabel 3.2 Jumlah Data Penelitian	42
Tabel 3.3 Contoh <i>Cleaning</i>	44
Tabel 3.4 Contoh <i>Case Folding</i>	45
Tabel 3.5 Contoh <i>Normalization</i>	46
Tabel 3.6 Contoh <i>Tokenizing</i>	46
Tabel 3.7 Contoh <i>Stopword Removal</i>	47
Tabel 3.8 Contoh <i>Stemming</i>	47
Tabel 3.9 Spesifikasi Perangkat Keras	53
Tabel 3.10 Perangkat Lunak	54
Tabel 4.1 Hasil Pengumpulan Data	55
Tabel 4.2 Jumlah Data pada Tahap Pembersihan	56
Tabel 4.3 Hasil <i>Cleaning</i>	57
Tabel 4.4 Hasil <i>Case Folding</i>	58
Tabel 4.5 Hasil <i>Normalization</i>	59
Tabel 4.6 Hasil <i>Tokenizing</i>	60
Tabel 4.7 Hasil <i>Stopword Removal</i>	61
Tabel 4.8 Hasil <i>Stemming</i>	62
Tabel 4.9 Hasil Pelabelan	63
Tabel 4.10 Hasil Analisis Sentimen.....	64
Tabel 4.11 Hasil Pembobotan TF-IDF	66
Tabel 4.12 Hasil <i>Train-Test Split</i>	66
Tabel 4.13 Hasil Random Oversampling.....	67
Tabel 4.14 Hasil Klasifikasi LinearSVC	68
Tabel 4.15 Hasil Klasifikasi Multinomial Naive Bayes	69
Tabel 4.16 Hasil Evaluasi LinearSVC dan Multinomial Naive Bayes	70

DAFTAR LAMPIRAN

<u>Lampiran 1 Repositori Kode Program Penelitian</u>	83
--	----