

**PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN
NAÏVE BAYES DALAM ANALISIS SENTIMEN DEMO DPR RI 2025**

PUBLIKASI ARTIKEL ILMIAH

Diajukan Untuk Memenuhi Salah Satu Syarat
Mencapai Gelar Sarjana komputer
Program Studi Informatika Universitas Sahid Surakarta



Disusun Oleh :

Lathif Nurma Huda
NIM 2022061013

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS, TEKNOLOGI DAN KESEHATAN
UNIVERSITAS SAHID SURAKARTA
2026**

LEMBAR PERSETUJUAN

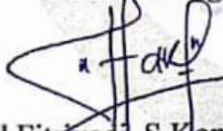
PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES* DALAM ANALISIS SENTIMEN DEMO DPR RI 2025

Disusun oleh :

Lathif Nurma Huda
NIM.2022061013

Publikasi Artikel Ilmiah ini telah disetujui untuk dipertahankan
Di hadapan dewan penguji
Pada tanggal 6 Februari 2026.

Pembimbing I



Farid Fitriyadi, S.Kom, M.Kom
NIDN.0607058902

Pembimbing II



Hardika Khusnuliawati, S.Kom, M.Kom
NIDN.0631089201

Mengetahui,

Ketua Program Studi



Hardika Khusnuliawati, S.kom, M.kom
NIDN.0631089201

LEMBAR PENGESAHAN

PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES* DALAM ANALISIS SENTIMEN DEMO DPR RI 2025

Disusun Oleh :

LATHIF NURMA HUDA
NIM.2022061013

Publikasi Artikel Ilmiah ini telah diterima dan disahkan

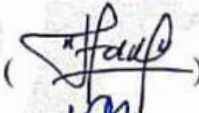
Oleh dewan penguji artikel ilmiah

Program Studi Informatika

Universitas Sahid Surakarta

Pada hari Jumat tanggal 13 Februari

Dewan Penguji

- | | | |
|----------------|--|---|
| 1. Penguji I | Farid Fitriyadi, S.Kom.,M.Kom
NIDN.0607058902 | () |
| 2. Penguji II | Hardika Khusnuliawati, S.Kom.,M.Kom
NIDN.0631089201 | () |
| 3. Penguji III | Dr.Sri Huning Anwariningsih, S.T, M.Kom
NIDN.0017067901 | () |

Mengetahui,

**Ketua Program Studi
Informatika**



Hardika Khusnuliawati, S.Kom.,M.Kom
NIDN. 0631089201

**Dekan Fakultas Sains, Teknologi
Dan Kesehatan**



Apt. Apwan, S.Farm., M.Sc
NIDN.0626088401

**SURAT PERNYATAAN
ORISINALITAS KARYA ILMIAH**

Saya mahasiswa Program Studi Informatika Fakultas Sains, Teknologi, Dan Kesehatan Universitas Sahid Surakarta yang bertanda tangan di bawah ini:

Nama : Lathif Nurma Huda

NIM : 2022061013

Menyatakan dengan sesungguhnya bahwa Publikasi Artikel Ilmiah :

Judul : PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*
DAN *NAÏVE BAYES* DALAM ANALISIS SENTIMEN DEMO DPR
RI 2025

Adalah benar-benar karya yang saya susun sendiri. Apabila dikemudian hari terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, saya bersedia menerima sanksi sesuai peraturan yang telah berlaku di Universitas Sahid Surakarta termasuk pencabutan gelar sarjana yang telah saya peroleh.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya dan apabila dikemudian hari terbukti melakukan kebohongan maka saya akan menanggung konsekuensinya.

Surakarta, Ferbruari 2026

akan,

(Lathif Nurma Huda)

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Sahid Surakarta, saya yang bertanda tangan di bawah ini:

Nama : Lathif Nurma Huda
NIM : 2022061013
Program Studi : Informatika
Fakultas : SAINS, TEKNOLOGI, DAN KESEHATAN
Email : hudalathif@gmail.com
Jenis Karya : Publikasi Artikel Ilmiah

Demikian perkembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Sahid Surakarta Hak Bebas Royalti Non Eksklusif (Non-Exclusive Royalty Free Right) atas skripsi saya yang berjudul:

PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE DAN NAÏVE BAYES DALAM ANALISIS SENTIMEN DEMO DPR RI 2025

Badan instrument/desain/perangkat (jika ada). Berhak menyimpan, mengalihkan bentuk, digitalisasi, mengolah dalam bentuk pangkalan data (database), merawat dan mempublikasikan karya ilmiah saya selama tetap mencantumkan nama saya sebagai penulis (author) dan pembimbing sebagai cauthor atau pencipta dan sebagai hak milik.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa paksaan dari pihak manapun.

Dibuat di: Surakarta

Pada Tanggal: Ferbruari 2026

nyataan,


(Lathif Nurma Huda)

MOTTO

“ Jika kamu bersyukur, maka Aku akan menambah nikmat kepadamu.”

(Q.S Ibrahim : 7)

“Hal besar akan datang kepada orang yang bersabar.”

(EF)

~ Time is gold, Real Madrid vs Man City UCL 2022.



KATA PENGANTAR

Alhamdulillahirobbilalamin, segala puji bagi Allah SWT karena limpahan nikmat dan rahmat-Nya, sehingga penulis dapat menyelesaikan Laporan Skripsi dengan judul “Perbandingan Algoritma *Support Vector Machine* dan *Naïve Bayes* dalam Analisis Sentimen Demo DPR RI 2025”. Proposal ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer Pada Fakultas Sains, Teknologi dan Kesehatan Universitas Sahid Surakarta. Penyusunan proposal ini terwujud berkat bantuan, arahan, bimbingan dan doa dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Ibu Dr.Sri Huning Anwariningsih, S.T, M.Kom. selaku Rektor Universitas Sahid Surakarta.
2. Bapak Apt. Ahwan, S.Farm., M.Sc. selaku Dekan Fakultas Sains, Teknologi dan Kesehatan Universitas Sahid Surakarta.
3. Bapak Farid Fitriyadi, S.Kom, M.Kom. selaku Dosen Pembimbing I.
4. Ibu Hardika Khusnuliawati, S.Kom, M.Kom. selaku Ketua Program Studi Informatika Universitas Sahid Surakarta dan selaku Dosen Pembimbing II.
5. Seluruh dosen program studi Informatika Universitas Sahid Surakarta yang telah mendukung dan memberikan ilmunya.
6. Bapak, ibu, dan keluarga yang telah memberikan doa dan dukungan.
7. Terima kasih secara khusus saya sampaikan kepada Effy Febriyanti, yang telah menjadi sumber semangat dan ketenangan di tengah lelahnya proses penyusunan artikel ini. Terima kasih telah menjadi pendengar yang baik bagi setiap keluh kesah saya dan selalu percaya pada kemampuan saya bahkan di saat saya meragu.
8. Teman-teman dan pihak lain yang tidak dapat penulis sebutkan satu persatu yang telah membantu penulis dalam proses penyusunan skripsi ini.

Besar harapan penulis laporan ini dapat berguna bagi pembaca dan Universitas Sahid Surakarta.

Surakarta, 2026

Lathif Nurma Huda

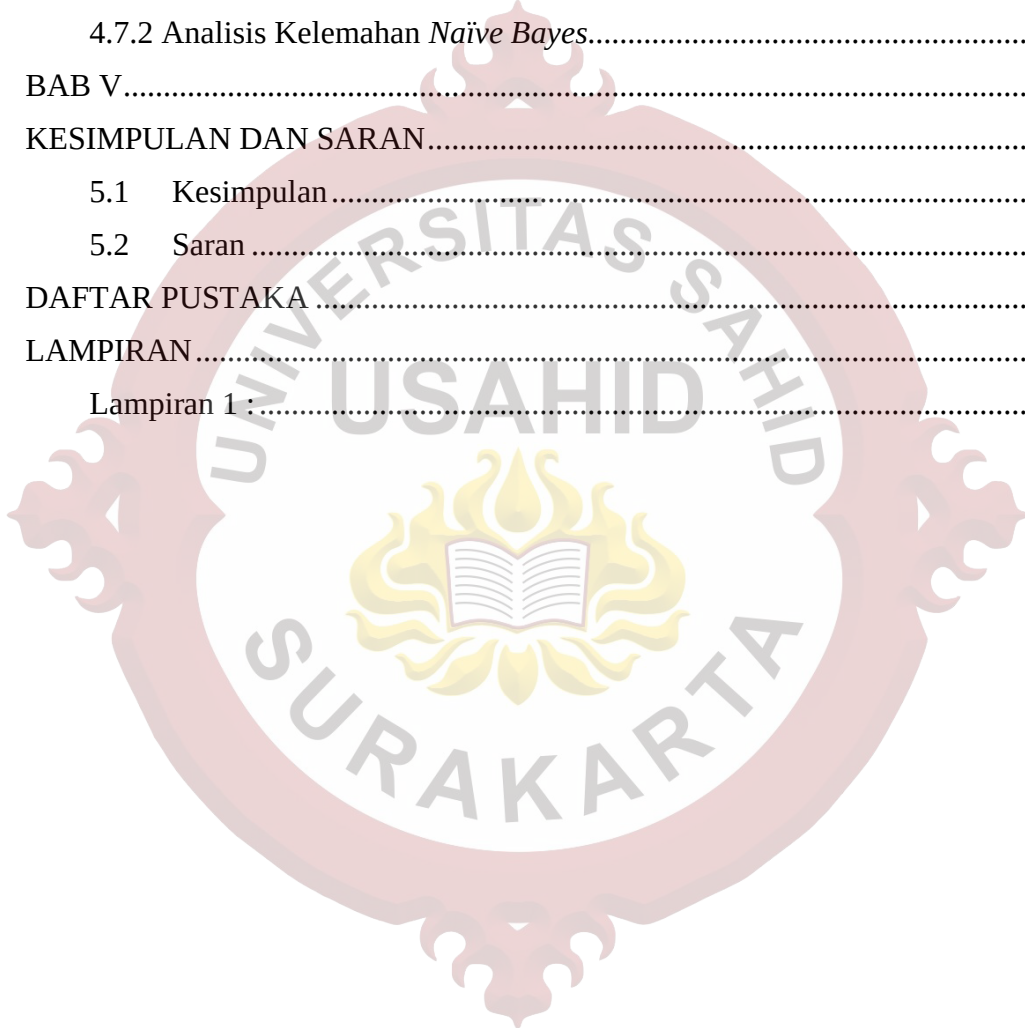


DAFTAR ISI

PUBLIKASI ARTIKEL ILMIAH	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
SURAT PERNYATAAN.....	iv
ORISINALITAS KARYA ILMIAH	iv
PERNYATAAN PERSETUJUAN PUBLIKASI	v
KARYA ILMIAH KEPENTINGAN AKADEMIS	v
MOTTO.....	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
ABSTRAK	xiv
ABSTRACT.....	xv
BAB I	1
PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan dan Manfaat Penelitian	4
1.4.1 Tujuan	4
1.4.2 Manfaat	4
BAB II.....	5
TINJAUAN PUSTAKA.....	5
2.1 Penelitian Terdahulu	5
2.2 Kerangka Pemikiran.....	8
2.3 Landasan Teori.....	13
2.3.1 Analisis Sentimen (<i>Sentiment Analysis</i>).....	13
2.3.2 Pra-pengolahan Data Text (<i>Text Preprocessing</i>)	13
2.3.3 Ekstraksi Fitur (<i>Bag-of-Word</i>)	14

2.3.3 Perlabelan Berbasis Leksikon (<i>Lexicon-Based</i>).....	15
2.3.4 Algoritma <i>Support Vector Machine</i>	15
2.3.5 Algoritma <i>Naïve Bayes</i>	16
2.3.6 Optimasi Model dengan <i>Grid Search Cross Validation</i>	16
2.3.7 K-Fold Cross Validation	17
2.3.8 Evaluasi Model.....	17
BAB III.....	19
METODOLOGI PENELITIAN.....	19
3.1 Sumber Data.....	19
3.2 Alat dan Bahan Penelitian.....	20
3.2.1 Perangkat Keras (<i>Hardware</i>)	20
3.2.1 Perangkat Lunak (<i>Software</i>).....	20
3.3 Alur Penelitian	21
3.3.1 Pengumpulan Data	21
3.3.2 Pra-Pengolahan Data (<i>Preprocessing</i>)	22
3.3.3 Perlabelan Data (<i>Data Labeling</i>)	23
3.3.4 Ekstraksi Fitur (<i>Feature Extraction</i>).....	23
3.3.5 Pembagian Data (<i>Data Splitting</i>)	23
3.3.6 Permodelan (<i>Modeling</i>).....	24
3.3.7 Optimasi Model (<i>Hyperparameter Tuning</i>).....	24
3.3.8 Evaluasi Model.....	25
BAB IV	26
HASIL DAN PEMBAHASAN.....	26
4.1 Deskripsi Data Penelitian.....	26
4.2 Hasil Pra-Pengolahan Data (<i>Preprocessing</i>).....	27
4.3 Hasil Perlabelan Data.....	29
4.4 Ekstraksi Fitur dan Pembagian Data	30
4.4.1 Ekstraksi Fitur (<i>Feature Selection</i>)	30
4.4.2 Pembagian Data (<i>Data Splitting</i>)	31
4.5 Validasi Label Data (Perhitungan Manual).....	31
4.5.1 Perhitungan Manual <i>Support Vector Machine</i>	31
4.5.2 Perhitungan Manual <i>Naïve Bayes</i>	33

4.5.3 Validasi Perhitungan Manual.....	35
4.6 Hasil Pengujian dan Evaluasi Model	36
4.6.1 Hasil Pengujian Parameter Standar	36
4.6.2 Hasil Pengujian Parameter Tuned (<i>Hyperparameter Tuned</i>).....	37
4.7 Pembahasan Hasil Penelitian	38
4.7.1 Analisis Keunggulan <i>Support Vector Machine</i> (SVM)	39
4.7.2 Analisis Kelemahan <i>Naïve Bayes</i>	39
BAB V.....	41
KESIMPULAN DAN SARAN.....	41
5.1 Kesimpulan.....	41
5.2 Saran	42
DAFTAR PUSTAKA	43
LAMPIRAN.....	46
Lampiran 1 :	46



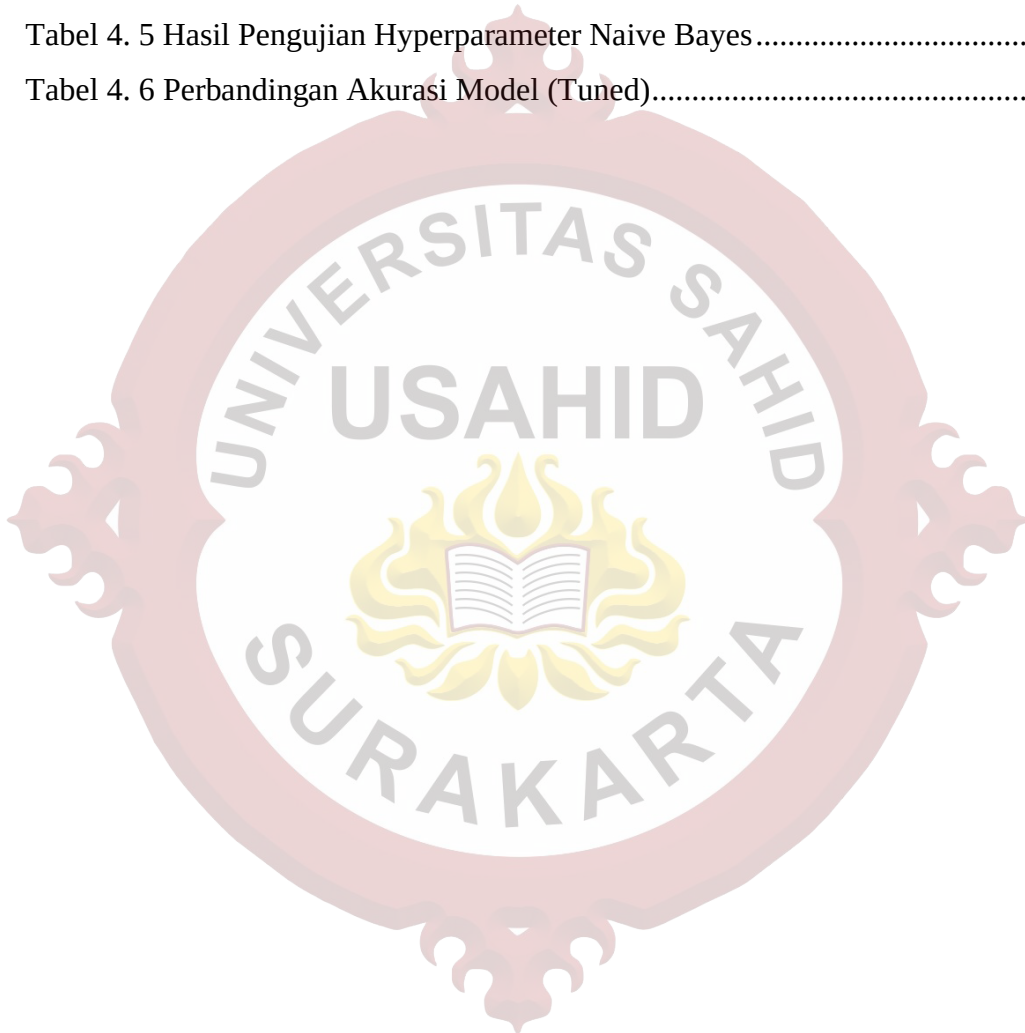
DAFTAR GAMBAR

Gambar 2. 1 Kerangka Pemikiran.....	9
Gambar 3. 1 Diagram Alur Penelitian.....	21
Gambar 4. 1 <i>Wordcloud</i> setelah pengolahan data.....	29
Gambar 4. 2 Visualisasi distribusi sentimen komentar YouTube.....	30
Gambar 4. 3 Kinerja Model (Parameter <i>Default</i>).....	37
Gambar 4. 4 Matriks Korelasi SVM yang Dioptimalkan	39
Gambar 4. 5 Matriks Korelasi <i>Naive Bayes</i> yang Dioptimalkan	40



DAFTAR TABEL

Tabel 4. 1 Contoh Proses Pra-Pengolahan Data.....	28
Tabel 4. 2 Distribusi Kelas Sentimen.....	30
Tabel 4. 3 Hasil Validasi Manual dan Sistem	35
Tabel 4. 4 Hasil Pengujian Hyperparameter SVM.....	37
Tabel 4. 5 Hasil Pengujian Hyperparameter Naive Bayes	38
Tabel 4. 6 Perbandingan Akurasi Model (Tuned).....	38



ABSTRAK

Media sosial YouTube menjadi platform utama bagi masyarakat untuk menyuarakan opini terkait peristiwa politik, seperti Demo DPR RI 2025. Volume komentar yang masif memerlukan analisis sentimen otomatis untuk memahami opini publik. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* (NB) dalam mengklasifikasikan sentimen komentar tersebut ke dalam tiga kelas: Positif, Negatif, dan Netral. Data dikumpulkan dari tiga video YouTube terkait demo menggunakan YouTube Data API v3, menghasilkan 18.973 komentar unik setelah pra-pengolahan yang dilabeli menggunakan metode Lexicon-Based (InSet). Fitur diekstraksi menggunakan *CountVectorizer* (*Bag-of-Words*) dan data dibagi dengan rasio 80:20. Optimasi model dilakukan melalui *Hyperparameter Tuning* dengan *GridSearchCV*. Hasil analisis menunjukkan dominasi sentimen Negatif sebesar 40,34%. Pada pengujian klasifikasi, SVM dengan parameter terbaik (*kernel* 'linear' dan $C=10$) berhasil mencapai akurasi tertinggi sebesar 90,17%. Kinerja ini secara signifikan mengungguli *Naive Bayes* ($\alpha=0.5$) yang hanya mencapai akurasi 66,14%. Kesimpulannya, SVM terbukti sebagai model yang lebih superior dan andal untuk klasifikasi sentimen pada dataset komentar politik berbahasa Indonesia yang berdimensi tinggi.

Kata Kunci : Analisis Sentimen, *Support Vector Machine*, *Naive Bayes*, Komentar YouTube, Demo DPR RI 2025.

ABSTRACT

YouTube serves as a primary social media platform for the public to voice opinions regarding political events, such as the 2025 DPR RI Demonstration. The massive volume of comments necessitates automated sentiment analysis to comprehend public opinion. This study aims to compare the performance of Support Vector Machine (SVM) and Naive Bayes (NB) algorithms in classifying comment sentiments into three classes: Positive, Negative, and Neutral. Data was collected from three demonstration-related YouTube videos using the YouTube Data API v3, resulting in 18,973 unique comments after preprocessing, which were labeled using the Lexicon-Based method (InSet). Features were extracted using CountVectorizer (Bag-of-Words), and the data was split into an 80:20 ratio. Model optimization was performed through Hyperparameter Tuning using GridSearchCV. Descriptive analysis results indicated a dominance of Negative sentiment at 40.34%. In classification testing, SVM with the best parameters ('linear' kernel and $C=10$) achieved the highest accuracy of 90.17%. This performance significantly outperformed Naive Bayes ($\alpha=0.5$), which only attained an accuracy of 66.14%. In conclusion, SVM proved to be a superior and reliable model for sentiment classification on high-dimensional Indonesian political comment datasets.

Keywords : Sentiment Analysis, Support Vector Machine, Naive Bayes, YouTube Comments, 2025 DPR RI Demonstration.

