



**ANALISIS SENTIMEN WARGANET YOUTUBE PADA VIDEO DEBAT  
CAPRES-CAWAPRES MENGGUNAKAN ALGORITMA  
NAIVE BAYES DAN K-NEAREST NEIGHBORS**

**LAPORAN TUGAS AKHIR**

**AKBAR Satria Wijaksono**  
**41518010089**

**UNIVERSITAS  
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2024**



**ANALISIS SENTIMEN WARGANET YOUTUBE PADA VIDEO DEBAT  
CAPRES-CAWAPRES MENGGUNAKAN ALGORITMA  
NAIVE BAYES DAN KNN**

**LAPORAN TUGAS AKHIR**

**AKBAR Satria Wijaksono**

**41518010089**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**UNIVERSITAS  
MERCU BUANA  
PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2024**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Akbar Satria Wijaksono  
NIM : 41518010089  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Warganet YouTube pada Video Capres-Cawapres Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbors.

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 30 Juli 2024



Akbar Satria Wijaksono

UNIVERSITAS  
MERCU BUANA

## HALAMAN PENGESAHAN

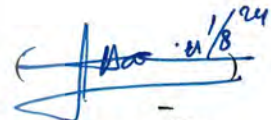
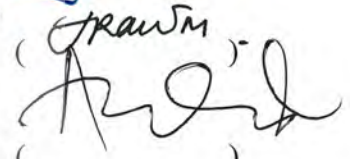

Laporan Skripsi ini diajukan oleh:

Nama : Akbar Satria Wijaksono  
NIM : 41518010089  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Warganet YouTube pada Video Capres-Cawapres Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbors.

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata Satu (S1) pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh :

Pembimbing : Lukman Hakim, ST, M.Kom  
NIDN : 0327107701  
Ketua Penguji : Ratna Mutu Manikam, S.Kom., MT  
NIDN : 308017101  
Penguji 1 : Anis Cherid, S.E., M.TI.  
NIDN : 0328127203  
Penguji 2 : Ida Farida, S.T., M.Kom.  
NIDN : 0324018301

  
( Lukman Hakim )  
  
( Ratna Mutu Manikam )  
  
( Anis Cherid )

Jakarta, 30 Juli 2024

Mengetahui,

Dekan

  
Dr. Bambang Jekonowo, S.Si., MTI  
NIDN : 0320037002

Ketua Program Studi

  
Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701

## KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridhanya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Lukman Hakim, ST, M.Kom selaku dosen pembimbing akademik.
5. Bapak Lukman Hakim, ST, M.Kom selaku dosen pembimbing MPTI yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
6. Bapak Misni, S.Kom., M.Kom selaku dosen pengampu.
7. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
8. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, Juli 2024

Akbar Satria Wijaksono



## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Akbar Satria Wijaksono  
NIM : 41523100100  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Warganet YouTube pada Video Capres-Cawapres Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbors.

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS  
MERCU BUANA

Jakarta, 30 Juli 2024

Yang menyatakan,



Akbar Satria Wijaksono

## ABSTRAK

Nama : Akbar Satria Wijaksono  
NIM : 41518010089  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Warganet YouTube pada Video Capres-Cawapres Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbors.  
Dosen Pembimbing : Lukman Hakim, ST, M.Kom

Abstrak - Pemilu atau pemilihan umum merupakan bagian dari dasar demokrasi suatu negara untuk memilih wakil rakyat dalam pemerintahan. Pemilihan wakil rakyat untuk menjadi presiden di tahun 2024 ini menjadi panas karena banyak menarik perhatian dengan perdebatan capres-cawapres yang dianggap dapat menentukan popularitas atau elektabilitas kandidat. Debat capres – cawapres dapat dilihat pada kanal YouTube yang bisa di analisis melalui komentar atau opini masyarakat yang terdapat pada video debat. Analisis sentimen masyarakat melalui platform Youtube menggunakan algoritma Naive Bayes dan K-Nearest Neighbors (KNN) ini memperlihatkan pengaruh opini warganet terhadap elektabilitas paslon kedalam bentuk sentimen positif, negatif dan netral. Sentimen ini dapat digunakan dalam melihat calon mana yang memiliki citra positif, negatif atau netral pada warganet pengguna Youtube. Hasil penelitian, algoritma Naive Bayes terbukti lebih unggul dibandingkan algoritma KNN dalam analisis sentimen komentar video YouTube tentang debat capres cawapres, dengan nilai akurasi masing-masing 0.74 dan 0.67. Ini menunjukkan bahwa Naive Bayes lebih efektif dalam menangkap pola dan tren dalam teks komentar, membuatnya menjadi pilihan yang lebih baik untuk digunakan dalam menganalisis data.

**Kata kunci:** Pemilu, Capres-Cawapres, Naive Bayes, K-Nearest Neighbors.

## ABSTRACT

Nama : Akbar Satria Wijaksono  
NIM : 41518010089  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Analisis Sentimen Warganet YouTube pada Video Capres-Cawapres Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbors.  
Dosen Pembimbing : Lukman Hakim, ST, M.Kom

*Abstract - General elections are part of the basis of a country's democracy to elect representatives of the people in government. The election of people's representatives to become president in 2024 is hot because it attracts a lot of attention with presidential and vice presidential debates which are considered to determine the popularity or electability of candidates. The presidential and vice presidential debates can be seen on the YouTube channel which can be analyzed through public comments or opinions contained in the debate video. Analysis of public sentiment through the Youtube platform using the Naive Bayes and K-Nearest Neighbors (KNN) algorithms shows the influence of netizen opinions on candidate electability in the form of positive, negative and neutral sentiments. This sentiment can be used to see which candidates have a positive, negative or neutral image on Youtube users. The results of the study, the Naive Bayes algorithm proved superior to the KNN algorithm in sentiment analysis of YouTube video comments on the presidential and vice presidential debates, with accuracy values of 0.74 and 0.67 respectively. This shows that Naive Bayes is more effective in capturing patterns and trends in the comment text, making it a better choice to use in data analysis.*

*Keywords: Election, presidential and vice presidential candidates, Naive Bayes, K-Nearest Neighbors.*



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>HALAMAN PERNYATAAN KARYA SENDIRI .....</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>iv</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS .....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>viii</b>
<b>DAFTAR TABEL .....</b>	<b>x</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xiii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	2
1.3 Tujuan Penelitian .....	2
1.4 Manfaat Penelitian .....	2
1.5 Batasan Masalah .....	3
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>4</b>
2.1 Penelitian Terdahulu .....	4
2.2 Teori Pendukung.....	11
<b>BAB III METODE PENELITIAN .....</b>	<b>14</b>
3.1 Jenis Penelitian.....	14
3.2 Tahapan Penelitian.....	14
<b>BAB IV PEMBAHASAN .....</b>	<b>18</b>
4.1 Dataset .....	18
4.2 Text Processing .....	33
4.3 Membuat Label Sentimen .....	41
4.4 Visualisasi Data dengan Analisis per Grafik .....	48
4.5 Pembobotan TF-IDF .....	56

4.6	Implementasi Algoritma Naive Bayes .....	57
4.7	Implementasi Algoritma K-Nearest Neighbors (KNN).....	60
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>63</b>
5.1	Kesimpulan .....	63
5.2	Saran .....	63
<b>DAFTAR PUSTAKA .....</b>		<b>64</b>
<b>LAMPIRAN.....</b>		<b>66</b>
Lampiran 1 Kartu Asistensi .....		66
Lampiran 2 Lampiran Halaman Pernyataan Luaran Tugas Akhir (Khusus Publikasi Ilmiah).....		67
Lampiran 3 Lampiran Naskah Artikel Jurnal .....		68
Lampiran 4 Curriculum Vitae .....		73
Lampiran 5 Surat Pernyataan HAKI.....		74
Lampiran 6 Sertifikat BNSP .....		76
Lampiran 8 Form Revisi Dosen Penguji.....		77
Lampiran 9 Hasil Cek Turnitin .....		79



## DAFTAR TABEL

Tabel 1 Perangkat Keras .....	6
Tabel 2 Perangkat Lunak .....	6



## DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian .....	14
Gambar 4.1 Jumlah Data.....	18
Gambar 4.2 Import Library .....	18
Gambar 4.3 Read/Load Dataset YouTube Channel Najwa Shihab .....	19
Gambar 4.4 Menampilkan jumlah baris dan kolom.....	19
Gambar 4.5 Nama Kolom DataFrame “data_najwa_shihab” .....	20
Gambar 4.6 Menampilkan data duplikat.....	20
Gambar 4.7 Missing value/nilai data yang hilang .....	20
Gambar 4.8 Isi DataFrame “data_najwa_shihab” .....	21
Gambar 4.9 Read/Load Dataset YouTube Channel KOMPASTV .....	22
Gambar 4.10 Menampilkan jumlah data “data_kompastv” .....	23
Gambar 4.11 Nama Kolom “data_kompastv” .....	23
Gambar 4.12 Menampilkan data duplikat “data_kompastv” .....	24
Gambar 4.13 Menampilkan nilai data yang hilang pada “data_kompastv”.....	24
Gambar 4.14 Menampilkan isi DataFrame “data_kompastv” .....	24
Gambar 4.15 Read/Load Dataset YouTube Channel KPU RI.....	26
Gambar 4.16 Menampilkan jumlah data “data_kpu_ri” .....	26
Gambar 4.17 Menampilkan nama kolom “data_kpu_ri” .....	27
Gambar 4.18 Menampilkan data duplikat DataFrame “data_kpu_ri” .....	27
Gambar 4.19 Menampilkan nilai data yang hilang DataFrame “data_kpu_ri”.....	28
Gambar 4.20 Menampilkan 10 data yang hilang dari “data_kpu_ri” .....	29
Gambar 4.21 Mentransformasikan nilai yang hilang “data_kpu_ri” .....	30
Gambar 4.22 Menampilkan isi DataFrame “data_kpu_ri” .....	31
Gambar 4.23 Menggabungkan tiga variabel DataFrame .....	32
Gambar 4.24 Membuat kolom baru pada “coments_debate_df” .....	33
Gambar 4.25 Menyimpan data.....	33
Gambar 4.26 Case Folding .....	34
Gambar 4.27 Removal Punctuations.....	34
Gambar 4.28 Tokenization.....	35
Gambar 4.29 Stopwords.....	35
Gambar 4.30 “text_comment_tokenization” .....	36
Gambar 4.31 Tampilan 12 data terbawah .....	37
Gambar 4.32 Kata setelah diproses Stopwords.....	39
Gambar 4.33 Menampilkan 7 data teratas dari data “comments_debate_df”.....	39
Gambar 4.34 Menampilkan hasil pemisahan DataFrame .....	43
Gambar 4.35 Menampilkan hasil jumlah kemunculan kata.....	49
Gambar 4.36 Grafik 10 kata yang paling sering muncul .....	50
Gambar 4.37 Nilai korelasi dengan grafik Heatmap.....	51
Gambar 4.38 Perbandingan jumlah kemunculan data terhadap tiga Channel YouTube .....	52
Gambar 4.39 Perbandingan ketiga Lanel Sentimen.....	54
Gambar 4.40 Kata yang paling sering digunakan pada seluruh Label Sentimen .	55
Gambar 4.41 Nilai bobot setiap kata.....	57
Gambar 4.42 Output hasil nilai terbaik .....	58
Gambar 4.43 Grafik Confusion Matrix Algoritma NB.....	59

Gambar 4.44 Hasil Prediksi dan Klasifikasi Algoritma NB .....	59
Gambar 4.45 Output nilai terbaik algoritma KNN .....	60
Gambar 4.46 Grafik Confusion Matrix Algoritma KNN.....	61
Gambar 4.47 Hasil prediksi dan klasifikasi algoritma KNN .....	61
Gambar 4.48 Kompari tingkatan nilai akurasi .....	62





## DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi .....	18
Lampiran 2 Lampiran Halaman Pernyataan Luaran Tugas Akhir (Khusus Publikasi Ilmiah).....	19
Lampiran 3 Lampiran Naskah Artikel Jurnal (Khusus Publikasi Ilmiah) .....	20
Lampiran 4 Curriculum Vitae .....	21
Lampiran 5 Surat Pernyataan HAKI.....	22
Lampiran 6 Sertifikat BNSP .....	24
Lampiran 7 Surat Ijin Riset Perusahaan.....	25

