

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
PHK DI INDONESIA PADA *TWITTER*
MENGUNAKAN *NAÏVE BAYES*
DAN *SUPPORT VECTOR*
MACHINE (SVM)**

SKRIPSI

**Oleh:
Abdu Malik AlHakim
202110715209**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

(DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Tugas akhir : Analisis Sentimen Masyarakat Terhadap PHK di
Indonesia Pada *Twitter* Menggunakan *Naïve*
Bayes Dan *Support Vector Machine* (SVM)
Nama Mahasiswa : Abdu Malik AlHakim
Nomor Pokok : 202110715209
Mahasiswa
Program Studi/Fakultas : Informatika/Ilmu Komputer
Tanggal Lulus Ujian : 07 Juli 2025
Skripsi

Jakarta, 10 Juli2025

MENYETUJUI,

Pembimbing I

Pembimbing II

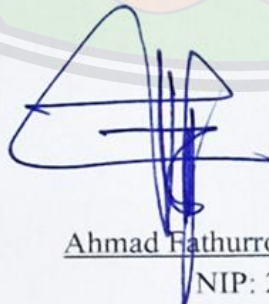


Prima Dina Atika, S.Kom., M.Kom.
NIDN: 0311037107



Dr. Herlawati, S.Si., M.M., M.Kom.
NIDN: 0311097302

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I.
NIP: 2012486

PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA

2025

LEMBAR PENGESAHAN

Judul Tugas akhir : ANALISIS SENTIMEN MASYARAKAT
TERHADAP PHK DI INDONESIA PADA
TWITTER MENGGUNAKAN *NAÏVE*
BAYES DAN *SUPPORT VECTOR*
MACHINE (SVM)
Nama Mahasiswa : Abdu Malik AlHakim
Nomor Pokok Mahasiswa : 202110715209
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 07 Juli 2025

Jakarta, 18 Juli 2025

MENGESAHKAN,

Ketua Tim Penguji : Rafika Sari, S.Si., M.Si.
NIDN : 0329098902

Penguji II : Ir. Muhammad Khaerudin, M.Kom
NIDN : 0413066604

Penguji III : Prima Dina Atika, S.Kom., M.Kom.
NIDN : 0311037107

MENGETAHUI,

Ketua
Program Studi Informatika



Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Dekan
Fakultas Ilmu Komputer



Dr. Robertus Suraji, S.S., M.A.
NIP. 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Abdu Malik AlHakim
NPM : 202110715209
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Masyarakat Terhadap PHK di Indonesia Pada *Twitter* Menggunakan *Naïve Bayes* dan *Support Vector Machine* (SVM)

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak mana pun.

Bekasi, 18 Juli 2025
Penulis



Abdu Malik AlHakim

ABSTRAK

Abdu Malik AlHakim . 202110715209 . analisis sentimen masyarakat terhadap phk di indonesia pada *twitter* menggunakan *naïve bayes* dan *support vector machine* (svm) . Bekasi: Fakultas Ilmu Komputer . Universitas Bhayangkara Jakarta Raya. 2025

Fenomena pemutusan hubungan kerja (PHK) di Indonesia telah memunculkan berbagai opini publik, khususnya di media sosial. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap isu PHK menggunakan data dari Twitter, serta membandingkan performa dua algoritma klasifikasi teks, yaitu *Naïve Bayes* dan *Support Vector Machine*. Pendekatan *Knowledge Discovery in Databases* digunakan sebagai kerangka kerja penelitian, yang mencakup tahapan pemilihan data, pembersihan teks, transformasi, klasifikasi, dan evaluasi. Sebanyak 3.458 *tweet* dikumpulkan dan diproses melalui tahapan pra-pemrosesan, kemudian diklasifikasikan ke dalam sentimen positif dan negatif. Penilaian performa dilakukan dengan tiga skenario pembagian data latih dan uji: 90:10, 80:20 dan 70:30. Hasil menunjukkan bahwa *Support Vector Machine* memberikan akurasi tertinggi sebesar 84,93% pada skenario 90:10, dibandingkan *Naïve Bayes* dengan akurasi 82,61% pada skenario yang sama. Visualisasi melalui *wordcloud* turut digunakan untuk memperkuat interpretasi terhadap kata-kata dominan dalam opini publik. Temuan ini menunjukkan bahwa algoritma klasifikasi dapat dimanfaatkan untuk memahami persepsi masyarakat terhadap isu ketenagakerjaan dan mendukung pengambilan keputusan berbasis data sosial. Penelitian ini dapat dikembangkan lebih lanjut dengan memperluas cakupan data dan mengevaluasi metode yang lebih kompleks untuk meningkatkan akurasi klasifikasi.

Kata Kunci : Sentimen; Pemutusan Hubungan Kerja (PHK); *Twitter*; *Naïve Bayes*; *Support Vector Machine* (SVM)

ABSTRACT

Abdu Malik AlHakim . 202110715209 . public sentiment analysis on layoffs in indonesia on twitter using naïve bayes and support vector machine (svm) . Bekasi: Faculty of Computer Science. Jakarta Bhayangkara University. 2025

The phenomenon of layoffs in Indonesia has led to various public opinions, especially on social media. This research aims to analyze public sentiment on the layoff issue using data from Twitter, and compare the performance of two text classification algorithms, namely Naïve Bayes and Support Vector Machine. The Knowledge Discovery in Databases approach is used as the research framework, which includes the stages of data selection, text cleaning, transformation, classification, and evaluation. A total of 3,458 tweets were collected and processed through the pre-processing stage, then classified into positive and negative sentiments. Performance assessment was conducted with three scenarios of training and test data sharing: 90:10, 80:20 and 70:30. The results showed that Support Vector Machine gave the highest accuracy of 84.93% in the 90:10 scenario, compared to Naïve Bayes with 82.61% accuracy in the same scenario. Visualization through wordcloud was also used to strengthen the interpretation of dominant words in public opinion. The findings show that classification algorithms can be utilized to understand public perceptions of employment issues and support social data-based decision-making. This research can be further developed by expanding data coverage and evaluating more complex methods to improve classification accuracy.

Keywords : Sentiment; Layoff; Twitter; Naïve Bayes; Support Vector Machine (SVM)

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala, atas segala rahmat, taufik, dan hidayah-nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul **"Analisis Sentimen Masyarakat Terhadap PHK di Indonesia Pada *Twitter* Menggunakan *Naïve Bayes* Dan *Support Vector Machine* (SVM)"**. Tugas akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.

Dalam penyusunan tugas akhir ini, penulis telah mendapatkan bantuan, bimbingan, serta dukungan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis sampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Irjen Pol (Purn) Dr. Drs. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku Dosen Pembimbing I, yang telah memberikan arahan, masukan, doa, dan dorongan selama proses penulisan tugas akhir ini.

5. Ibu Dr. Herlawati, S.Si., M.M., M.Kom. selaku Dosen Pembimbing II, yang telah memberikan koreksi, saran, doa, dan dorongan selama proses penulisan tugas akhir ini.
6. Bapak Tri Dharma Putra, S.T., M.Sc. selaku Dosen Pembimbing Akademik, atas bimbingan, nasihat, dan arahan yang telah diberikan selama masa studi hingga penyusunan tugas akhir ini.
7. Kepada seluruh Dosen Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya, atas ilmu, wawasan yang telah diberikan selama penulis menempuh pendidikan.
8. Kepada Ayah dan Ibu tercinta, terima kasih atas cinta, doa, dan dukungan tanpa syarat yang telah mengiringi setiap langkah penulis hingga tugas akhir ini dapat terselesaikan.
9. Kepada Widi Lestari yang menemani langkah demi langkah, menjadi tempat istirahat sejenak dalam hiruk pikuk penyusunan tugas akhir, penulis menyampaikan rasa terima kasih atas kehadiran dan ketulusannya.
10. Terima kasih kepada Fadjar, Shatria, Raihan, Syehan, Hakim, Hikmal, dan rekan seperjuangan atas kebersamaan, motivasi, dan kerja sama yang turut mewarnai perjalanan studi penulis.

Bekasi, Juni 2025

Hormat saya

Abdu Malik AlHakim

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	5
1.4 Batasan Masalah.....	6
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian	7
1.6.1 Manfaat Teoritis	7
1.6.2 Manfaat Praktis	7
1.7 Sistematika Penulisan	8
BAB II TINJAUAN PUSTAKA	10
2.1 Penelitian Terdahulu.....	10

2.1.1	<i>Tabel Related Research / State of The Art</i>	15
2.2	Teori Pendukung	20
2.2.1	<i>Artificial Intelligence (AI)</i>	20
2.2.2	<i>Machine Learning</i>	21
2.2.3	<i>Data Mining</i>	22
2.2.4	<i>Web Scraping</i>	24
2.2.5	<i>Text Mining</i>	26
2.2.6	<i>Text Preprocessing</i>	27
2.2.7	Media Sosial.....	29
2.2.8	<i>Twitter</i>	30
2.2.9	Pemutusan Hubungan Kerja (PHK)	31
2.2.10	<i>Knowledge Discovery in Databases (KDD)</i>	32
2.2.11	Analisis Sentimen	34
2.2.12	<i>Term Frequency – Invers Document Frequency (TF-IDF)</i>	35
2.2.13	Algoritma <i>Naïve Bayes</i>	36
2.2.14	Algoritma <i>Support Vector Machine (SVM)</i>	40
2.2.15	<i>Confusion Matrix</i>	42
2.2.16	<i>Python</i>	45
BAB III METODOLOGI PENELITIAN		47
3.1	Objek Penelitian	47
3.2	Kerangka Penelitian	48
3.2.1	<i>Data Selection (Pemilihan Data)</i>	49
3.2.2	<i>Preprocessing (Pra-proses Data)</i>	50
3.2.3	<i>Transformation (Transformasi Data)</i>	52
3.2.4	<i>Data Mining (Penambangan Data)</i>	53

3.2.5	<i>Evaluation</i> (Evaluasi).....	55
3.2.6	<i>Knowledge</i> (Pengetahuan).....	56
3.3	Analisis Kebutuhan <i>Hardware</i> dan <i>Software</i>	56
3.3.1	Spesifikasi Perangkat Keras	56
3.3.2	Spesifikasi Perangkat Lunak	57
BAB IV HASIL DAN PEMBAHASAN.....		58
4.1	<i>Data Selection</i> (Pemilihan Data).....	58
4.2	<i>Preprocessing</i> (Pra-proses Data).....	60
4.2.1	<i>Case Folding</i>	60
4.2.2	<i>Cleansing</i>	61
4.2.3	<i>Tokenization</i>	63
4.2.4	<i>Normalization</i>	64
4.2.5	<i>Remove Stopword</i>	65
4.2.6	<i>Stemming</i>	67
4.2.7	<i>Labeling</i>	69
4.3	<i>Transformation</i> (Tranformasi Data).....	72
4.3.1	<i>Splitting Data</i> (Pembagian Data)	72
4.3.2	Pembobotan TF-IDF (<i>Term Frequency – Invers Document Frequency</i>)	75
4.4	<i>Data Mining</i> (Penambangan Data).....	78
4.4.1	<i>Naïve Bayes</i>	78
4.4.2	<i>Support Vector Machine</i> (SVM).....	83
4.5	<i>Evaluation</i> (Evaluasi).....	86
4.5.1	Model <i>Naïve Bayes</i>	86
4.5.2	Model <i>Support Vector Machine</i> (SVM)	93

4.6	<i>Knowledge</i> (Pengetahuan).....	100
4.6.1	Model Klasifikasi Terbaik.....	101
4.6.2	<i>Wordcloud</i>	102
BAB V PENUTUP		109
5.1	Kesimpulan	109
5.2	Saran.....	110
DAFTAR PUSTAKA.....		112
LAMPIRAN.....		117



DAFTAR TABEL

Tabel 2.1 <i>Related Research / State of The Art</i>	15
Tabel 2.2 Tabel <i>Confusion Matrix</i>	43
Tabel 4.1 Hasil penerapan <i>Case Folding</i>	61
Tabel 4.2 Hasil tahapan <i>Cleansing</i>	62
Tabel 4.3 Hasil penerapan <i>Tokenization</i>	63
Tabel 4.4 Hasil penerapan <i>Normalization</i>	64
Tabel 4.5 Hasil tahap <i>Remove Stopword</i>	66
Tabel 4.6 Hasil penerapan <i>Stemming</i>	68
Tabel 4.7 Contoh data untuk prediksi manual <i>Naïve Bayes</i>	79
Tabel 4.8 Frekuensi kata D5 di data latih.....	81
Tabel 4.9 Contoh data untuk prediksi manual SVM.....	84
Tabel 4.10 Contoh TF-IDF untuk perhitungan SVM.....	84
Tabel 4.11 Hasil evaluasi model <i>Naïve Bayes</i> dan SVM	99
Tabel 4.12 Waktu komputasi model <i>Naïve Bayes</i> dan SVM.....	100

DAFTAR GAMBAR

Gambar 1.1 Data Jumlah PHK di Indonesia (2020-2024)	1
Gambar 2.1 Tahapan <i>Knowledge Discovery in Databases</i> (KDD)	33
Gambar 2.2 Algoritma SVM menemukan <i>hyperplane</i>	40
Gambar 2.3 <i>Small margin</i> dan <i>Large Margin</i> terbentuk oleh SVM	41
Gambar 3.1 Tampilan halaman utama <i>Twitter</i>	47
Gambar 3.2 Kerangka Penelitian	48
Gambar 3.3 Contoh tampilan <i>tweet</i> topik PHK	49
Gambar 3.4 <i>Dataset</i> hasil <i>scraping tweet</i>	50
Gambar 3.5 Tahap <i>Transformation</i> (Trasformasi Data)	52
Gambar 3.6 Proses Algoritma <i>Naïve Bayes</i>	53
Gambar 3.7 Proses Algoritma <i>Support Vector Macine</i> (SVM)	54
Gambar 4.1 Contoh Kode <i>scraping</i> data dari <i>twitter</i>	59
Gambar 4.2 <i>Dataset</i> mentah hasil <i>scraping</i>	59
Gambar 4.3 Pemilihan kolom <i>full_text</i>	60
Gambar 4.4 Kode <i>Case folding</i>	61
Gambar 4.5 Kode untuk tahapan <i>Cleansing</i>	62
Gambar 4.6 Kode untuk <i>Tokenization</i>	63
Gambar 4.7 Kode untuk <i>Normalization</i>	64
Gambar 4.8 Kode untuk <i>Remove Stopword</i>	66
Gambar 4.9 Kode untuk <i>Stemming</i>	67
Gambar 4.10 Kode untuk <i>Labeling otomatis</i>	69
Gambar 4.11 Contoh penerapan <i>Labeling</i> otomatis	69

Gambar 4.12 Distribusi sentimen hasil <i>Labeling</i>	70
Gambar 4.13 Diagram <i>pie</i> distribusi sentimen.....	70
Gambar 4.14 Hasil penerapan <i>Label Encoder</i>	71
Gambar 4.15 Kode untuk perbandingan 90:10	73
Gambar 4.16 Hasil penerapan <i>splitting data</i> 90:10.....	73
Gambar 4.17 Kode untuk perbandingan 80:20	74
Gambar 4.18 Hasil penerapan <i>splitting data</i> 80:20.....	74
Gambar 4.19 Kode untuk perbandingan 70:30	75
Gambar 4.20 Hasil penerapan <i>splitting data</i> 70:30.....	75
Gambar 4.21 Kode untuk pembobotan TF-IDF	76
Gambar 4.22 Hasil penerapan TF-IDF.....	77
Gambar 4.23 Fitur dengan bobot TF-IDF tertinggi	77
Gambar 4.24 Kode untuk menerapkan <i>Multinomial Naive Bayes</i>	79
Gambar 4.25 Kode untuk menerapkan <i>Support Vector Machine</i> (SVM)	83
Gambar 4.26 <i>Confusion Matrix Naive Bayes</i> dengan skenario 90:10	87
Gambar 4.27 <i>Confusion Matrix Naive Bayes</i> dengan skenario 80:20	89
Gambar 4.28 <i>Confusion Matrix Naive Bayes</i> dengan skenario 70:30	91
Gambar 4.29 <i>Confusion Matrix SVM</i> dengan skenario 90:10.....	93
Gambar 4.30 <i>Confusion Matrix SVM</i> dengan skenario 80:20.....	95
Gambar 4.31 <i>Confusion Matrix SVM</i> dengan skenario 70:30.....	97
Gambar 4.32 Perbandingan nilai akurasi	101
Gambar 4.33 Kode untuk menampilkan <i>Wordcloud Naive Bayes</i>	102
Gambar 4.34 <i>Wordcloud</i> positif <i>Naive Bayes</i>	103
Gambar 4.35 <i>Wordcloud</i> negatif <i>Naive Bayes</i>	104
Gambar 4.36 Kode untuk menampilkan <i>Wordcloud SVM</i>	105

Gambar 4.37 <i>Wordcloud</i> positif SVM.....	105
Gambar 4.38 <i>Wordcloud</i> negatif SVM.....	107



DAFTAR LAMPIRAN

Lampiran 1 : Uji Plagiarisme	117
Lampiran 2 : Biodata Peneliti	118
Lampiran 3 : Kartu Konsultasi Bimbingan	119
Lampiran 4 : Rekomendasi dari pembimbing untuk Tugas Akhir	122
Lampiran 5 : Publikasi Artikel	123

