

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

**SENTIMEN ANALISIS TERHADAP PILKADA JAKARTA 2024
BERDASARKAN PLATFORM MEDIA SOSIAL TWITTER (X)
MENGUNAKAN MACHINE LEARNING**



SKRIPSI

**Diajukan Sebagai Syarat Untuk Menyelesaikan Pendidikan
Program Sarjana (S-1)
Pada Program Studi Sistem Informasi**

Oleh:

BIANKHA ARIESTY

NIM : 2102030039

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU TEKNIK

UNIVERSITAS BINA INSAN

2024

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

HALAMAN PENGESAHAN SKRIPSI



**SENTIMEN ANALISIS TERHADAP PILKADA JAKARTA 2024
BERDASARKAN PLATFORM MEDIA SOSIAL TWITTER (X)
MENGUNAKAN *MACHINE LEARNING***

Oleh :

BIANKHA ARIESTY

NIM : 2102030039

Lubuklinggau, Januari 2025

Pembimbing I

Pembimbing II

Muhammad Nur Alamsyah, M.Kom

Nelly Khairani Daulay, M.Kom

**Mengesahkan
Dekan Fakultas Ilmu Teknik
Universitas Bina Insan,**

(Dr. Rudi Kurniawan, S.T., M.Kom)

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

HALAMAN PERSETUJUAN TIM PENGUJI SKRIPSI



Pada hari Jum'at tanggal 24 bulan Juli tahun 2025 telah dilaksanakan sidang skripsi oleh Program Studi Sistem Informasi Fakultas Ilmu Teknik Universitas Bina Insan.

Nama : Biankha Ariesty
NIM : 2102030039
Judul Skripsi : Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan *Machine Learning*

Komisi Penguji

1. Ketua : **Muhammad Nur Alamsyah, M.Kom** (.....)
2. Sekretaris : **Nelly Khairani Daulay, M.Kom** (.....)
3. Anggota : **Dr. Rudi Kurniawan, S.T., M.Kom** (.....)

Mengetahui
Kepala Program Studi Sistem Informasi
Fakultas Ilmu Teknik
Universitas Bina Insan

(Harma Oktafia Lingga Wijaya, M.Kom)

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

HALAMAN MOTTO DAN PERSEMBAHAN



Motto:

- ❖ *“Cause there were pages turned with the bridges burned everything you lose is a step you take. So, make the friendship bracelets, take the moment and taste it You’ve got no reason to be afraid. You’re on your own kid. Yeah, You can face this”-Taylor Swift*
- ❖ *Sekali nyebur ya harus basah, sekali berjuang ya harus dituntaskan. Jika kamu menyerah, untuk apa kamu memulai.*
- ❖ *Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. QS. Al-Baqarah:286.*

Persembahan Kepada:

- ❖ *Ibu dan Bapak tercinta, yang telah memberikan dukungan dan doa untuk kelancaran semua urusanku*
- ❖ *Diriku sendiri yang aku cintai*
- ❖ *Cicik tercinta yang selalu membantuku*
- ❖ *Kakak tercinta dan Dua kucing kesayanganku (kibo & kiyo), yang selalu bisa mengembalikan moodku*
- ❖ *Teman-teman seperjuangan*
- ❖ *Almamaterku*

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

HALAMAN PERNYATAAN



Saya yang bertanda tangan di bawah ini :

Nama Mahasiswa/i : Biankha Ariesty

NIM : 2102030039

Program Studi : Sistem Informasi

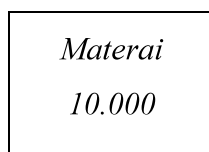
Fakultas : Ilmu Teknik

Menyatakan dengan sesungguhnya bahwa penelitian dan penulisan Skripsi yang saya susun sebagai persyaratan untuk memperoleh gelar Sarjana (S-1) Universitas Bina Insan, merupakan hasil kerja saya sendiri dan tidak menyuruh orang lain yang mengerjakannya. Ada pun bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain dan telah saya tuliskan sumbernya secara jelas sesuai dengan norma, kaidah dan etika penulisan ilmiah.

Jika dikemudian hari ternyata terbukti bahwa penelitian dan tugas akhir ini bukan hasil kerja saya sendiri atau plagiat dalam bagian-bagian tertentu, maka saya bersedia dikenakan sanksi sesuai dengan peraturan perundangan yang berlaku.

Lubuklinggau, Januari 2025

Penulis



Biankha Ariesty

NIM. 2102030039

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

ABSTRACT



In each country, there are various processes for selecting state or regional leaders. In Indonesia itself there are 2 processes for selecting leaders, namely Pemilu (General Election) and Pilkada (Regional Head Election). In this research, the data taken are opinions uploaded via the Twitter platform (X). Tweets often express clear opinions, which helps the algorithm more easily identify text sentiment as positive, negative or neutral. This research will compare six algorithms, namely Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Artificial Neural Network, Random Forest and Decision Tree in sentiment classification. The data taken using the keywords for the 2024 Jakarta regional election was 26,354 with a time period of January 2024-January 2025. The test results show that Support Vector Machine (SVM) is a better algorithm with higher performance. The precision value for negative sentiment is 85%. Neutral sentiment is 83% and positive sentiment is 95%. The recall value for negative sentiment is 80%, neutral sentiment is 91%, positive sentiment is 93%. The accuracy value reaches 90%, and the ROC AUC value is 97% and the average K-10 Cross Validation value is 90%.

Keywords: 2024 Jakarta Regional Election, sentiment analysis, twitter, machine learning

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

ABSTRAK



Di setiap negara, terdapat berbagai macam proses pemilihan pemimpin negara ataupun daerah. Di Indonesia sendiri memiliki 2 proses pemilihan pemimpin, yaitu Pemilu (Pemilihan Umum) dan Pilkada (Pemilihan Kepala Daerah). Pada penelitian ini, data yang diambil adalah Opini yang diunggah melalui platform twitter (X), Tweet sering kali mengungkapkan pendapat yang jelas, yang membantu algoritma lebih mudah mengidentifikasi sentimen teks sebagai positif, negatif, atau netral. Penelitian ini akan membandingkan enam algoritma, yaitu *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* dalam klasifikasi sentimen. Data yang diambil menggunakan kata kunci pilkada Jakarta 2024 sebanyak 26.354 dengan rentang waktu Januari 2024-Januari 2025. Hasil pengujian menunjukkan *Support Vector Machine* (SVM) menjadi algoritma yang lebih baik dengan kinerja yang lebih tinggi. Nilai *precision* untuk sentimen negatif sebesar 85% Sentimen netral sebesar 83% dan sentimen positif sebesar 95%. Pada nilai *recall* untuk sentimen negatif adalah 80%, sentimen netral adalah 91%, sentimen positif adalah 93%. Untuk nilai *accuracy* mencapai 90%, dan nilai ROC AUC sebesar 97% serta nilai rata-rata *Cross Validation K-10* sebesar 90%.

Kata Kunci: Pilkada Jakarta 2024, sentimen analisis, twitter, *machine learning*

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

KATA PENGANTAR

Puji syukur penulis puji atas kehadiran Allah SWT yang telah memberikan limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi penelitian dengan judul "Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan Machine Learning" dengan baik.

Skripsi ini bertujuan untuk memberikan pemahaman yang komprehensif tentang tren sentimen publik dan unsur-unsur yang membentuk opini publik dalam ranah politik digital. Dalam penyusunan, penulis menyadari bahwa skripsi ini tidak terlepas dari bimbingan, dukungan, dan kontribusi dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Allah SWT yang telah memberikan Kesehatan dan kenikmatan, sehingga proposal skripsi dapat terselesaikan dengan baik
2. Kedua Orang Tua Penulis, Ibu Sunarti dan Bapak Syamsu Eppendi yang telah memberikan banyak bantuan, dukungan dan doa kepada penulis sehingga skripsi ini dapat selesai dengan baik
3. Rektor Universitas Bina Insan, Bapak Dr. H. Sardiyo, M.M
4. Wakil Rektor I Universitas Bina Insan, Bapak Muhammad Akbar, S.T.,M.IT
5. Wakil Rektor II Universitas Bina Insan, Bapak Wakhid Nur Mukhlis, M.Pd.,M.M
6. Dekan Fakultas Ilmu Teknik Universitas Bina Insan, Bapak Dr. Rudi Kurniawan, S.T.,M.Kom yang telah memberikan arahan dan bimbingan dalam penulisan skripsi ini
7. Kepala Program Studi Sistem Informasi Universitas Bina Insan, Ibu Harma Oktafia Lingga Wijaya, M.Kom yang telah memberikan bimbingan dan arahan
8. Bapak Muhammad Nur Alamsyah, M.Kom selaku pembimbing I yang banyak memberikan arahan mengenai penulisan skripsi
9. Ibu Nelly Khairani Daulay, M.Kom selaku pembimbing II yang banyak memberikan bimbingan dalam kepenulisan skripsi

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

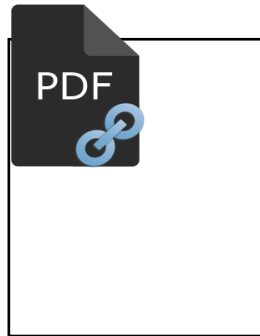
10. Seluruh staf, karyawan dan dosen Universitas yang telah banyak membantu penulis
11. Teman yang kebersamaan penulis dari masa SMA-Sekarang, Rizka Nur Aziza dan Rita Purnama yang telah banyak menerima keluhan kesah, memberi semangat kepada penulis
12. Kakak dan dua kucing tersayang yang selalu mengembalikan mood penulis dengan tingkah lucunya.
13. Seluruh teman-teman prodi Sistem Informasi Angkatan 2021 yang tidak dapat disebutkan satu persatu

Penulis berharap penelitian ini akan memberikan kontribusi yang signifikan terhadap kemajuan ilmiah, khususnya dalam bidang teknologi informasi, analisis data, dan sosial-politik. Diharapkan bahwa proposal ini akan menjadi titik awal yang berharga bagi para peneliti lain dan pihak yang berkepentingan.

Lubuklinggau, 25 Januari 2025

Penulis

Protected by PDF Anti-Copy Free
(Upgrade to Pro Version to Remove the Watermark)
DAFTAR RIWAYAT HIDUP



Biodata

Nama : Biankha Ariesty
Tempat/Tanggal Lahir : Lahat, 14 April 2004
Jenis Kelamin : Perempuan
Agama : Islam
Alamat : Jl. Air Temam

Pendidikan

- SD : SDN 58 Lubuklinggau
- SMP : SMPN 9 Lubuklinggau
- SMA : MAN 2 Lubuklinggau

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

DAFTAR ISI

	Halaman
Halaman Judul	i
Halaman Pengesahan	ii
Halaman Persetujuan	iii
Halaman Motto dan Persembahan	iv
Halaman Pernyataan	v
<i>Abstract</i>	vi
Abstrak	vii
Kata Pengantar	viii
Daftar Riwayat Hidup	x
Daftar Isi	xi
Daftar Tabel	xiii
Daftar Gambar	xiv
Daftar Lampiran	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penelitian	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Batasan Masalah	3
1.5 Tujuan dan Manfaat Penelitian	4
1.5.1 Tujuan Penelitian	4
1.5.2 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II KAJIAN PUSTAKA	6
2.1 Literatur	6
2.1.1 Sentimen Analisis	6
2.1.2 Pilkada Jakarta 2024	6
2.1.3 Platform Media Sosial Twitter (X)	6
2.1.4 <i>Naïve Bayes</i>	6
2.1.5 <i>Support Vector Machine</i>	7
2.1.6 <i>K-Nearest Neighbor</i>	7
2.1.7 <i>Artificial Neural Network</i>	7
2.1.8 <i>Random Forest</i>	7
2.1.9 <i>Decision Tree</i>	8
2.1.10 <i>Machine Learning</i>	8
2.1.11 <i>Text Mining</i>	9
2.1.12 <i>Python</i>	9

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

2.1.13 <i>Preprocessing</i>	9
2.1.14 <i>CountVector</i>	10
2.1.15 Klasifikasi	10
2.1.16 <i>K-Fold Cross</i>	11
2.1.17 <i>Confusion M</i>	11
2.1.18 ROC AUC	11
2.2 Penelitian Relevan	12
2.3 Kerangka Berpikir	14
BAB III METODOLOGI PENELITIAN	16
3.1 Metode Penelitian	16
3.2 Metode Pengumpulan Data	16
3.2.1 Studi Pustaka	16
3.2.2 Data Primer	16
3.3 Metode Analisa	17
3.4 Tempat dan Waktu Penelitian	17
3.4.1 Tempat	17
3.4.2 Waktu	17
3.5 Alat dan Bahan	17
3.5.1 Alat	17
3.5.2 Bahan	18
3.6 Metode Pengujian dan Pengolahan Data	18
3.6.1 Metode Pengujian	18
3.6.2 Metode Pengolahan Data	20
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	27
4.1 Gambaran Umum	27
4.2 Hasil	27
4.2.1 Pengumpulan Data	27
4.2.2 <i>Preprocessing</i>	28
4.2.3 <i>Labelling</i>	31
4.2.4 Visualisasi	32
4.2.5 <i>CountVectorizer</i>	34
4.2.6 Klasifikasi	34
4.2.7 Evaluasi	34
4.3 Pembahasan	45
BAB V KESIMPULAN DAN SARAN	48
5.1 Kesimpulan	48
5.2 Saran	49
DAFTAR PUSTAKA	50
LAMPIRAN	54

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

DAFTAR TABEL




Halaman

Tabel 3. 1. Waktu Penelitian.....	17
Tabel 3.2. Confusion Matrix	19
Tabel 3.3. Contoh Hasil Cleansing.....	21
Tabel 3.4. Contoh Hasil Case Folding	21
Tabel 3.5. Contoh Hasil Tokenizing.....	22
Tabel 3. 6. Contoh Hasil Stopword Removal.....	23
Tabel 3. 7. Contoh Hasil Normalisasi	23
Tabel 3. 8. Contoh Hasil Translate to English.....	24
Tabel 3. 9. Contoh Hasil Stemming	24
Tabel 4. 1. Data Setelah <i>Preprocessing</i>	45
Tabel 4. 2. Label Sentimen Vader	46
Tabel 4. 3. Hasil Klasifikasi SVM	46
Tabel 4. 4. Hasil Klasifikasi <i>Naive Bayes</i>	46
Tabel 4. 5. Hasil Klasifikasi KNN	46
Tabel 4. 6. Hasil Klasifikasi ANN.....	47
Tabel 4. 7. Hasil Klasifikasi <i>Random Forest</i>	47
Tabel 4. 8. Hasil Klasifikasi <i>Decision Tree</i>	47

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

DAFTAR GAMBAR

	Halaman
 Gambar 2.1. Kerangka Berpikir.....	15
Gambar 3.1. Tahapan <i>Preprocessing</i>	20
Gambar 3.2. Tahapan <i>Labeling</i>	25
Gambar 3.3. Tahapan <i>CountVectorizer</i> dan Klasifikasi.....	26
Gambar 4.1. Hasil <i>Crawling</i>	28
Gambar 4.2. Hasil <i>Cleansing</i>	28
Gambar 4. 3. Hasil <i>Case Folding</i>	29
Gambar 4.4. Hasil <i>Tokenizing</i>	29
Gambar 4.5. Hasil <i>Stopword Removal</i>	30
Gambar 4.6. Hasil Normalisasi.....	30
Gambar 4.7. Hasil <i>Translate to English</i>	31
Gambar 4.8. Hasil <i>Stemming</i>	31
Gambar 4.9. Hasil <i>Labelling</i>	32
Gambar 4. 10. Hasil Komposisi Sentimen.....	32
Gambar 4. 11. Hasil Diagram Sentimen.....	33
Gambar 4. 12. Hasil <i>Wordcloud</i>	33
Gambar 4. 13. Hasil Evaluasi <i>Cross Validation K-10 SVM</i>	35
Gambar 4. 14. Hasil Evaluasi <i>Confusion Matrix SVM</i>	35
Gambar 4. 15. Hasil <i>Classification Report SVM</i>	36
Gambar 4. 16. Grafik ROC AUC SVM.....	36
Gambar 4. 17. Hasil Evaluasi <i>Cross Validation K-10 Naive Bayes</i>	37
Gambar 4. 18. Hasil Evaluasi <i>Confusion Matrix Naive Bayes</i>	37
Gambar 4. 19. Hasil <i>Classification Report Naive Bayes</i>	38
Gambar 4. 20. Grafik ROC AUC <i>Naive Bayes</i>	38
Gambar 4. 21. Hasil Evaluasi <i>Cross Validation K-10 KNN</i>	39
Gambar 4. 22. Hasil Evaluasi <i>Confusion Matrix KNN</i>	39
Gambar 4. 23. Hasil <i>Classification Report KNN</i>	39
Gambar 4. 24. Grafik ROC AUC KNN.....	40
Gambar 4. 25. Hasil Evaluasi <i>Cross Validation K-10 ANN</i>	40
Gambar 4. 26. Hasil Evaluasi <i>Confusion Matrix ANN</i>	41
Gambar 4. 27. Hasil <i>Classification Report ANN</i>	41
Gambar 4. 28. Grafik ROC AUC ANN.....	41
Gambar 4. 29. Hasil Evaluasi <i>Cross Validation K-10 Random Forest</i>	42
Gambar 4. 30. Hasil Evaluasi <i>Confusion Matrix Random Forest</i>	42
Gambar 4. 31. Hasil <i>Classification Report Random Forest</i>	43
Gambar 4. 32. Grafik ROC AUC <i>Random Forest</i>	43
Gambar 4. 33. Hasil Evaluasi <i>Cross Validation K-10 Decision Tree</i>	44
Gambar 4. 34. Hasil Evaluasi <i>Confusion Matrix Decision Tree</i>	44
Gambar 4. 35. Hasil <i>Classification Report Decision Tree</i>	44

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Gambar 4. 36. Grafik ROC AUC *Decision Tree*45



Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

DAFTAR LAMPIRAN



Lampiran 1. Form Pengajuan Ju.....	55
Lampiran 2. SK Pembimbing.....	56
Lampiran 3. SK Penguji.....	57
Lampiran 4. Dataset Awal.....	58
Lampiran 5. Dataset Setelah Preprocessing.....	59
Lampiran 6. Dataset Setelah Tahap Pelabelan.....	60
Lampiran 7. Lembar Bimbingan P1.....	61
Lampiran 8. Lembar Bimbingan P2.....	62
Lampiran 9. Lembar Perbaikan Ujian Skripsi.....	63

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

BAB I



1.1 Latar Belakang Penelitian

Di setiap negara, terdapat berbagai macam proses pemilihan pemimpin negara ataupun daerah. Di Indonesia sendiri memiliki 2 proses pemilihan pemimpin, yaitu Pemilu (Pemilihan Umum) terdiri dari pemilihan Presiden dan Wakil Presiden, DPR RI (Dewan Perwakilan Rakyat Republik Indonesia), DPRD (Dewan Perwakilan Rakyat Daerah) Provinsi, DPRD Kabupaten/Kota dan DPD RI (Dewan Perwakilan Daerah Republik Indonesia). Adapun proses pemilihan yang kedua yaitu Pilkada (Pemilihan Kepala Daerah) terdiri dari pemilihan Gubernur dan Wakil Gubernur serta Wali Kota dan Wakil Walikota. Proses tersebut dilaksanakan setiap 5 tahun sekali dan selalu menjadi topik hangat yang diperbincangkan oleh Masyarakat Indonesia.

Pembahasan tersebut sering dibahas pada platform sosial media, seperti facebook, Instagram, tiktok dan twitter (X). Komentar yang diberikan beragam, ada yang bersifat positif dan ada yang bersifat negatif. Pilkada Jakarta selalu menjadi topik hangat yang diperbincangkan, karena Jakarta merupakan pusat pemerintahan, ekonomi dan kehidupan sosial di Indonesia. Pemimpin yang terpilih dalam pemilihan kepala daerah akan memberikan pengaruh signifikan tidak hanya kepada warga Jakarta tetapi juga kepada kebijakan di tingkat nasional.

Algoritma *Naïve Bayes* merupakan metode yang digunakan untuk proses klasifikasi menggunakan statistik dan probabilitas [2]. Selanjutnya, algoritma *Support Vector Machine (SVM)* dapat menyelesaikan masalah klasifikasi dan regresi linier maupun nonlinier [3]. Algoritma *K-Nearest Neighbor (KNN)* merupakan klasifikasi data berdasarkan letaknya (jarak) dan kedekatannya dengan data lain [4]. Algoritma *Artificial Neural Network (ANN)* merupakan model yang meniru cara kerja jaringan neural biologis dirancang dan dilatih

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

untuk mempunyai kemampuan seperti manusia [5]. Algoritma Random Forest adalah sebuah algoritma yang menggunakan metode pemisahan biner secara berulang untuk mencapai titik optimal dalam struktur pohonnya, yang didasarkan pada pohon klasifikasi dan regresi [6]. Sedangkan algoritma Decision Tree adalah suatu model untuk klasifikasi dan prediksi dengan memanfaatkan sebuah struktur hierarkis yang berbentuk pohon [7].

Tujuan dari penelitian ini adalah untuk mengatasi masalah yang berkaitan dengan akurasi dan efektivitas algoritma machine learning dalam analisis sentimen berdasarkan data Twitter. Dengan mengevaluasi enam algoritma yang berbeda, tujuannya adalah untuk mengidentifikasi model yang paling efektif yang dapat mencerminkan sentimen publik secara menyeluruh dan akurat, khususnya terkait dengan Pilkada Jakarta 2024.

Pada penelitian ini, data yang diambil adalah Opini yang diunggah melalui platform twitter (X) yang bersifat negatif atau sentimen ini akan menjadi bahan analisis mengenai Pilkada Jakarta 2024. Tweet sering kali mengungkapkan pendapat yang jelas, yang membantu algoritma lebih mudah mengidentifikasi sentimen teks sebagai positif, negatif, atau netral. Analisis sentimen adalah proses yang dimulai dengan ekstraksi dan pemrosesan data dan diakhiri dengan memahami data yang pada dasarnya tidak beraturan dalam bentuk tekstual, dengan tujuan untuk memperoleh informasi dari suatu pendapat atau sekumpulan pendapat [1]. Analisis sentimen terhadap opini publik di media sosial berperan penting untuk memahami dinamika sosial dan politik dalam Masyarakat. Hasil analisis yang akurat memungkinkan semua pemangku kebijakan, baik Lembaga pemerintah, kandidat, maupun Masyarakat untuk membuat Keputusan yang lebih tepat yang menjunjung tinggi integritas demokrasi dan mempromosikan keharmonisan sosial.

Penelitian ini akan membandingkan enam algoritma, yaitu *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* untuk mengetahui lebih dalam tentang efektivitas algoritma tersebut dalam melakukan klasifikasi sentimen pada data twitter (X) terkait Pilkada Jakarta 2024.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Berdasarkan latar belakang di atas, penulis tertarik untuk melakukan penelitian dengan judul “*Seleksi Model Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan Machine Learning*”



1.2 Identifikasi Masalah

Dalam konteks Pilkada Jakarta 2024, penting untuk menganalisis opini Masyarakat. Oleh karena itu, penelitian ini diperlukan untuk mengidentifikasi performa terbaik antara algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest*, dan *Decision Tree* dalam mengklasifikasikan opini publik. Penelitian ini diharapkan memberikan wawasan mengenai pendapat Masyarakat pada Pilkada Jakarta 2024.

1.3 Rumusan Masalah

Adapun rumusan masalah pada penelitian ini, sebagai berikut :

- a. Bagaimana cara mengumpulkan data tweet mengenai Pilkada Jakarta 2024 pada sosial media X untuk analisis sentimen?
- b. Bagaimana proses preprocessing data tweet sebelum dilakukan klasifikasi sentimen?
- c. Bagaimana mengimplementasikan algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* dalam klasifikasi sentimen opini publik mengenai Pilkada Jakarta 2024 pada sosial media X?

1.4 Batasan Masalah

Karena luasnya ruang lingkup penelitian, maka fokus penelitian sebagai berikut:

- a. Data yang digunakan berupa tweet berbahasa Indonesia dengan rentang waktu Januari 2024 – Januari 2025 dengan kata kunci “pilkada jakarta 2024”.
- b. Klasifikasi sentimen menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

1.5 Tujuan dan Manfaat Penelitian

1.5.1 Tujuan Penelitian

1.5.1.1 Tujuan Umum

Tujuan umum dari penelitian ini untuk mengetahui opini publik terhadap Pilkada Jakarta 2024 pada sosial media X menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*.

1.5.1.2 Tujuan Khusus

Tujuan khusus dari penelitian ini sebagai syarat untuk menyelesaikan Pendidikan program sarjana (S-1) pada program studi sistem informasi di Universitas Bina Insan Kota Lubuklinggau. Selain itu, untuk mengukur Tingkat sentimen positif, negatif dan netral terhadap topik mengenai Pilkada Jakarta 2024 pada periode waktu tertentu untuk melihat persepsi publik.

1.5.2 Manfaat Penelitian

1.5.2.1 Manfaat Bagi Ilmu Pengetahuan

Sebagai pengembangan ilmu data science khususnya di bidang analisis sentimen menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* dalam menganalisis opini publik mengenai Pilkada Jakarta 2024.

1.5.2.2 Manfaat Bagi Umum

Sebagai Gambaran opini publik mengenai Pilkada Jakarta 2024, terutama kepada kandidat, pemerintah dan Masyarakat mengenai pada media sosial X.

1.5.2.3 Manfaat Bagi Peneliti

Menerapkan ilmu pengetahuan yang telah didapatkan untuk mengembangkan kemampuan riset dalam analisis data, khususnya pengolahan data teks tidak terstruktur dari media sosial.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

1.6 Sistematika Penulisan

Dalam penulisan skripsi merupakan laporan dari hasil penelitian, direncanakan terdiri dari masing-masing bab berisi:

BAB I : PENDAHULUAN

Dalam bab ini terdapat latar belakang penelitian, rumusan masalah, Batasan masalah, Tujuan dan Manfaat penelitian serta sistematika penulisan.

BAB II : KAJIAN PUSTAKA

Dalam bab ini terdapat teori-teori yang mendasari masalah yang diteliti

BAB III : METODOLOGI PENELITIAN

Dalam bab ini terdapat metode penelitian, metode pengumpulan data, metode Analisa, tempat dan waktu penelitian, alat dan bahan, metode pengujian dan pengolahan data.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini terdapat hasil penelitian dan pembahasan

BAB V : KESIMPULAN DAN SARAN

Bab ini terdapat Kesimpulan dari seluruh tahapan penelitian, serta berisikan saran ataupun masukan yang berguna bagi penelitian selanjutnya.

DAFTAR PUSTAKA

LAMPIRAN



2.1 Literatur

2.1.1 Sentimen Analisis

Analisis sentimen adalah proses yang dimulai dengan ekstraksi dan pemrosesan data dan diakhiri dengan memahami data yang pada dasarnya tidak beraturan dalam bentuk tekstual, dengan tujuan untuk memperoleh informasi dari suatu pendapat atau sekumpulan pendapat [1].

2.1.2 Pilkada Jakarta 2024

Pilkada Jakarta 2024 merupakan proses pemilihan kepala daerah yang dilaksanakan tidak hanya di Jakarta, tapi juga dilaksanakan di seluruh wilayah Indonesia setiap 5 tahun sekali. Dalam pemilihan kepala daerah tersebut terdapat pemilihan Gubernur dan Wakil Gubernur, Wali Kota dan Wakil Wali Kota.

2.1.3 Platform Media Sosial Twitter (X)

Platform Media Sosial Twitter (X) merupakan salah satu platform terbesar yang aktif digunakan oleh masyarakat Indonesia. Pada platform tersebut, pengguna dapat menggunakan 280 karakter untuk berinteraksi dengan pengguna lainnya, termasuk memberikan opini terhadap isu-isu yang sedang ramai diperbincangkan.

2.1.4 *Naïve Bayes*

Naïve Bayes merupakan metode yang digunakan untuk proses klasifikasi menggunakan statistik dan probabilitas. *Naïve Bayes* dikemukakan oleh ilmuwan dari Inggris, yaitu Thomas Bayes. Pada proses pengklasifikasiannya, *Naïve Bayes* memprediksi dengan

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

memanfaatkan pengalaman di masa lalu sehingga dikenal *Teorema Bayes* [2].



2.1.5 Support Vector Machine

Algoritma *Support Vector Machine (SVM)* dapat menyelesaikan masalah klasifikasi dan regresi linier maupun *nonlinier*. SVM juga digunakan untuk klasifikasi, misalnya *Support Vector Classification* dan regresi, seperti *Support Vector Regression*. Dalam pemodelan klasifikasi, SVM memiliki konsep yang lebih canggih dan lebih jelas secara matematis dibandingkan metode klasifikasi lainnya [3].

2.1.6 K-Nearest Neighbor

K-Nearest Neighbor (KNN) adalah teknik pembelajaran mesin yang digunakan dalam proses klasifikasi. Prinsip kerja KNN adalah mengklasifikasikan data berdasarkan kedekatan posisinya (jarak) dengan data lainnya [4]. Prinsip metode KNN sederhana, yaitu bekerja berdasarkan jarak terpendek dari sampel uji ke sampel pelatihan [8].

2.1.7 Artificial Neural Network

Algoritma Artificial Neural Network (ANN) adalah pemodelan berdasarkan otak manusia yang terdiri dari serangkain prosesor yang saling berhubungan dan sangat sederhana, disebut neuron. *Neuron* yang dihubungkan oleh bobot meneruskan sinyal dari suatu *neuron* ke *neuron* lainnya. Setiap neuron dapat memiliki banyak masukan dan satu keluaran. Jalur masukan ke suatu *neuron* dapat berupa data mentah atau data yang diproses oleh *neuron* sebelumnya. Sebaliknya, keluaran neuron dapat berupa hasil akhir ataupun bahan masukan untuk neuron berikutnya [5].

2.1.8 Random Forest

Random Forest (RF) adalah sebuah algoritma yang menggunakan metode pemisahan biner secara berulang untuk mencapai

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

titik akhir dalam struktur pohonnya, yang didasarkan pada pohon klasifikasi dan regresi. RF menghasilkan sejumlah pohon yang tidak tergantung satu sama lain. Setiap pohon memilih subset secara acak melalui bootstrap dari sampel data dan dari variabel input di setiap nodenya. Untuk proses klasifikasinya, RF mengadopsi pendekatan ansambel dari berbagai pohon dengan cara melihat kemunculan mayoritas untuk mencapai keputusan akhir [6].

2.1.9 Decision Tree

Decision Tree adalah algoritma yang termasuk dalam metode klasifikasi dalam data mining. Decision tree memiliki bentuk seperti diagram alir dengan susunan pohon, di mana setiap simpul mewakili atribut dan cabangnya menunjukkan hasil dari uji atau nilai dari atribut yang dimasukkan, sementara daun menggambarkan kelas atau pembagian kelas [9]. Proses klasifikasi ini memanfaatkan pengamatan pada simpul untuk menentukan target pada daun. Decision Tree adalah salah satu metode klasifikasi yang paling banyak digunakan karena mudah dipahami oleh manusia, serta kemampuannya dalam menyederhanakan proses pengambilan keputusan yang rumit menjadi lebih jelas dan mudah [10].

2.1.10 Machine Learning

Machine Learning (ML) atau pembelajaran mesin merupakan suatu pendekatan dalam AI yang sering digunakan untuk menggantikan atau meniru perilaku manusia dalam memecahkan masalah. *Machine Learning* setidaknya memiliki dua aplikasi utama, yaitu klasifikasi dan prediksi. Klasifikasi adalah metode yang digunakan untuk mengurutkan objek berdasarkan karakteristik tertentu. Sedangkan prediksi atau regresi digunakan untuk menyimpulkan keluaran dari data masukan berdasarkan data yang diperiksa oleh selama pelatihan [11].

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

2.1.11 *Text Mining*

Text mining adalah teknik untuk mengekstraksi informasi dari sekumpulan data tidak terstruktur berkualitas tinggi dan memperoleh data tentang suatu masalah tekstual pada topik tertentu. Teks mining memungkinkan untuk mengeksplorasi informasi penting dari sumber data [12].

2.1.12 *Python*

Python adalah Bahasa pemrograman berbasis objek yang interaktif. *Python* dirancang khusus untuk membantu programmer untuk membuat aplikasi mandiri ataupun pemrograman skrip dengan menghemat waktu, mudah dikembangkan dan kompatibel dengan sistemnya [13].

2.1.13 *Preprocessing*

Preprocessing bertujuan untuk mengubah data yang sebelumnya tidak terstruktur menjadi terstruktur [1]. Adapun tahap preprocessing diantaranya:

1. *Cleansing*

Merupakan tahapan pembersihan data dengan menghapus tanda baca dan karakter non-teks, serta menghapus data yang duplikat [14].

2. *Case Folding*

Merupakan tahapan mengubah huruf besar menjadi huruf kecil dengan menghilangkan semua tanda baca dalam kalimat [3].

3. *Tokenizing*

Dilanjutkan dengan tahapan *Tokenizing*, yang digunakan untuk memisahkan kata per kata [14].

4. Normalisasi

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Merupakan proses perubahan kata slang yang tidak mengikuti kata baku sesuai dengan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI) [14].

5. *Stopword Removal*

Pada tahap ini, dilakukan penghapusan kata-kata yang dianggap tidak penting pada saat proses klasifikasi. Hal ini dilakukan dengan menghilangkan kata yang tercantum dalam stoplist seperti kata dan, atau, dengan, di, dari, yang, dll [2].

6. *Stemming*

Pada tahap ini imbuhan dihilangkan atau bentuk kata diubah kembali menjadi baku [2].

2.1.14 *Count Vectorizer*

CountVectorizer adalah cara menghitung jumlah munculnya setiap kata dalam dokumen, memberikan representasi vektor yang menunjukkan angka munculnya setiap kata [15]. Perhitungan kemunculan kata merupakan metode yang mudah, tetapi jika kata-kata sering muncul dalam banyak dokumen, kemungkinan angka kemunculan kata yang tinggi mungkin tidak memiliki arti yang besar dalam vektor kepentingan [16].

2.1.15 *Klasifikasi*

Klasifikasi adalah evaluasi suatu objek data ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Klasifikasi membuat model berdasarkan data pelatihan yang ada. Kemudian menggunakan model untuk mengklasifikasikan data yang baru. Klasifikasi merupakan metode supervised learning, metode yang mencoba menemukan hubungan antara atribut masukan dan atribut target. Akurasi model klasifikasi dapat dilihat dari *accuracy*, *sensitivity*, *recall* dan *precision* yang didapatkan dari *confusion matrix* [11].

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

2.1.16 K-Fold Cross Validation

K-Fold Cross Validation adalah jenis uji validasi silang yang dirancang untuk mengukur kinerja proses metode algoritma dengan membagi sampel data secara acak dan mengelompokkan data berdasarkan nilai pemisahan K pada K-Fold. Kemudian salah satu kelompok k-fold digunakan sebagai data uji dan kelompok sisanya digunakan sebagai data latih [17].

2.1.17 Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi model dan ditampilkan secara visual dalam format tabel. Hal ini memungkinkan data scientist mendapatkan wawasan yang lebih mendalam tentang performa, kesalahan dan kelemahan model. Hal ini memungkinkan untuk melakukan penyesuaian dan menganalisis model lebih lanjut [18]. Perhitungan confusion matrix dalam menentukan accuracy, precision, recall dapat dilihat sebagai berikut:

1. Accuracy

Accuracy menunjukkan tingkat keakuratan klasifikasi model. Hal ini didefinisikan seberapa baik data pengujian disiapkan atau seberapa dekat nilai prediksi dengan nilai sebenarnya untuk mengukur performa model [19].

2. Precision

Precision merupakan kecepatan sistem memberikan hasil dalam menanggapi permintaan informasi pengguna [19].

3. Recall

Recall dapat didefinisikan sebagai ukuran jumlah dokumen teks penting [19].

2.1.18 ROC AUC

ROC (*Receiver Operating Characteristics*) adalah kurva yang menggambarkan seberapa baik kinerja model klasifikasi, yang

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

menampilkan efektivitas dua dimensi dari setiap klasifikasi yang diuji.

Sumbu horizontal menunjukkan jumlah *false positive*, sedangkan sumbu vertikal menunjukkan jumlah *true positive*. AUC (*Area Under Curve*) mengacu pada area di bawah grafik ROC. Kualitas klasifikasi dapat diklasifikasikan menurut nilai AUC-nya [20].

2.2 Penelitian Relevan

Dalam melakukan penelitian ini, dibutuhkan sumber yang relevan seperti jurnal yang dapat dijadikan sebagai acuan dengan penelitian terkait. Berikut merupakan penelitian relevan yang sudah dilakukan oleh peneliti sebelumnya:

Pertama, penelitian yang dilakukan oleh Eka Bayu Satriawan dkk tentang Analisis Sentimen Terhadap Pendapat Masyarakat Mengenai Pilkada 2024 Menggunakan Metode Support Vector Machine (SVM). Data pada penelitian diambil pada media sosial Twitter (X) dengan rentang waktu 3 Mei 2024-31 Mei 2024 menghasilkan data sebanyak 2.230 komentar. Hasil pengujian Support Vector Machine (SVM) berhasil mendapatkan performa tinggi dengan tingkat akurasi sekitar 93,09%, precision sekitar 93,13%, recall sekitar 93.09%, dan f1-score sekitar 93,08% [1].

Penelitian kedua dilakukan oleh Rizki Amanda dkk mengenai Analisis Sentimen Masyarakat Pada Teks Postingan dan Komentar di Platform X dalam Pemilihan Presiden 2024 di Indonesia Melalui Pemrograman Python dan Metode Naïve Bayes. Data pada penelitian diambil melalui media sosial Twitter (X) dengan total 156 tweet tentang calon presiden RI. Didapatkan hasil klasifikasi sentimen menggunakan metode Naive Bayes menghasilkan nilai akurasi sebesar 62,5 % atau 0,625 [21].

Penelitian ketiga dilakukan oleh Aluisius Dwiki Adhi Putra dan Safitri Juanita dengan judul Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit dan Bareksa dengan Algoritma KNN. Data pada penelitian diambil melalui ulasan pengguna aplikasi Bibit dan Bareksa dari google play store. Jumlah ulasan yang digunakan pada penelitian ini sebanyak 998 dengan rentang waktu 13 februari-20 April 2024. Berdasarkan hasil yang diperoleh dari tahapan

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

modelling dengan menggunakan algoritma K-Nearest Neighbor dan perbandingan 60:40 untuk training dan data testing, maka nilai akurasi precision dan recall yang di dapat dari tiap aplikasi yaitu untuk bibit 85,14%, 91,91%, dan 76,44% sedangkan untuk bareksa yaitu 81,70% , 87,15%, 75,73% [8].

Penelitian selanjutnya, dilakukan oleh David Galih Pradana dkk mengenai Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network. Penelitian ini mencakup pengukuran performa (akurasi, presisi, recall dan f-score) metode ANN dengan 304 data pasien penyakit jantung yang diperoleh dari pusat dataset Kaggle. Hasil dari pengukuran performa diperoleh nilai akurasi 73,77%, presisi 80,43%, recall 84,09% dan f1-score sebesar 82,22% [5].

Berikutnya, penelitian yang dilakukan oleh Rizki Wahyudi dan Gilang Kusumawardhana mengenai Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. Penelitian ini menggunakan lebih dari 1.000 review pengguna yang dikumpulkan dari aplikasi Grab Indonesia di Google play store. Hasil dari analisis menggunakan Support Vector Machine menghasilkan akurasi 85,54% dan Hasil Review positif yang paling sering diulas adalah "ovo", sedangkan review negatif yang paling sering diulas adalah "driver" [22].

Penelitian keenam dilakukan oleh Risha Nur Mauliza dan Yoannes Romando Sipayung mengenai Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes menggunakan data sebanyak 300 ulasan yang dibagi menjadi 2 kategori yaitu 100 data latih dan 200 data uji yang kemudian dilakukan pengujian menggunakan metode naïve bayes. Hasil pengujian data dengan metode naïve bayes didapati hasil dengan perolehan 103 sentimen positif, 47 sentimen negatif dan 50 sentimen netral [23].

Selanjutnya, penelitian yang dilakukan oleh Aldi Tangkelayuk dan Evangs Mailoa mengenai Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes Dan Decision Tree dengan menggunakan 2.081 data yang sudah bersih

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

menunjukkan metode K-nearest Neighbors memiliki tingkat akurasi paling tinggi sebesar 86.88% dibandingkan dengan Decision Tree sebesar 80.84% dan Naïve Bayes sebesar 78.54% sehingga metode K-nearest Neighbors merupakan metode yang paling akurat untuk klasifikasi data [24].

Penelitian berikutnya dilakukan oleh Dhita Diana Dewi dkk dengan judul Perbandingan Metode Neural Network Dan Support Vector Machine dalam Klasifikasi Diagnosa Penyakit Diabetes yang menggunakan 768 data. Berdasarkan hasil analisis pada penelitian ini didapatkan nilai akurasi untuk metode ANN sebesar 77,60%, sedangkan nilai akurasi untuk metode SVM sebesar 65,24%. Artinya penggunaan metode ANN lebih baik daripada SVM untuk mengklasifikasikan seseorang menderita diabetes atau tidak. Sedangkan untuk metode ANN memiliki nilai AUC sebesar 0,834 sehingga dapat dikategorikan dalam good classification [25].

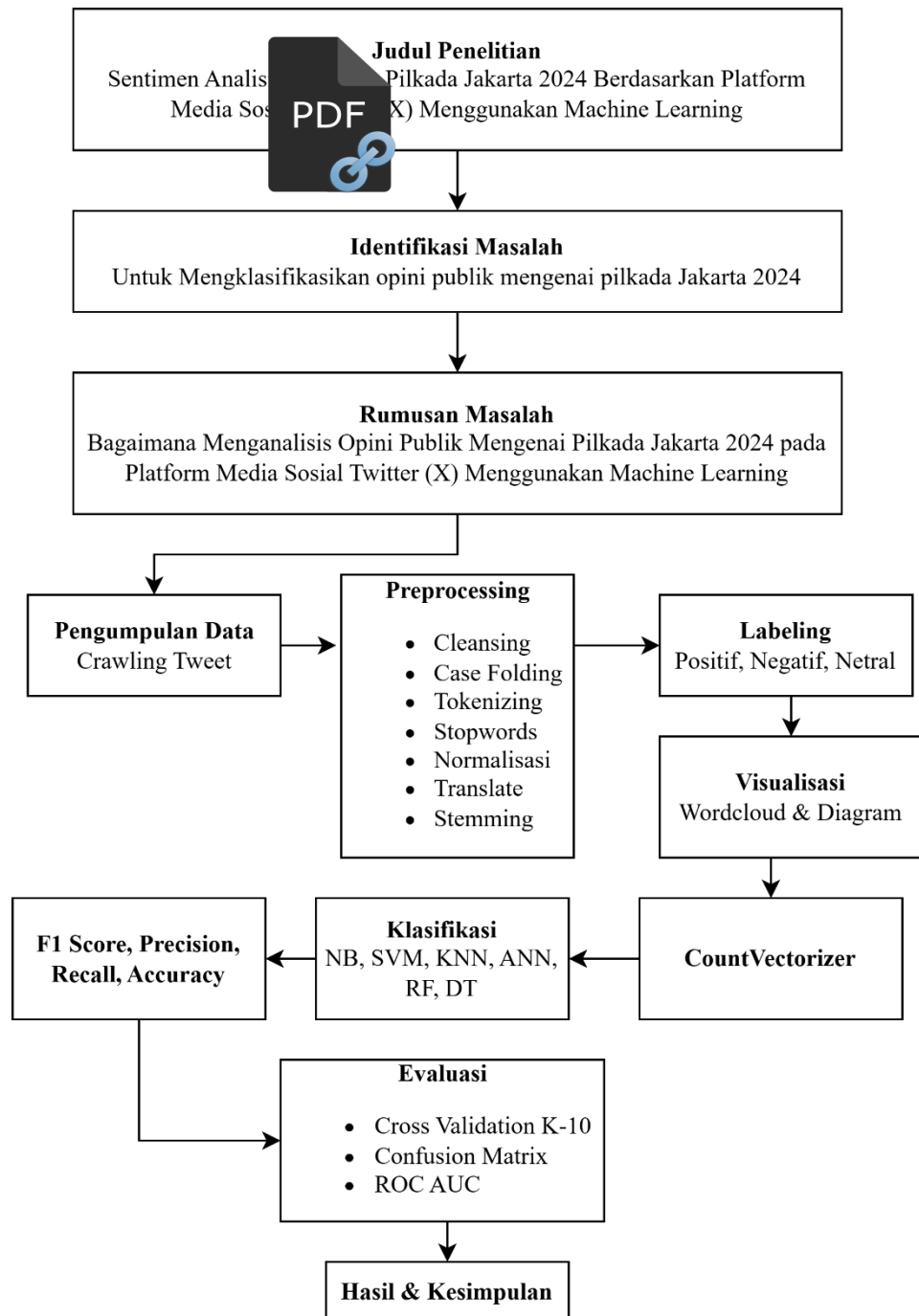
Berdasarkan penelitian relevan yang sudah ada, dapat disimpulkan bahwa algoritma Support Vector Machine memiliki akurasi paling tinggi dibandingkan dengan algoritma lainnya. Oleh karena itu, penulis tertarik melakukan klasifikasi menggunakan algoritma Support Vector Machine, Naïve Bayes, K-Nearest Neighbor dan Artificial Neural Network untuk mengkomparasi keempat algoritma dalam menganalisis sentimen Pilkada Jakarta 2024.

2.3 Kerangka Berpikir

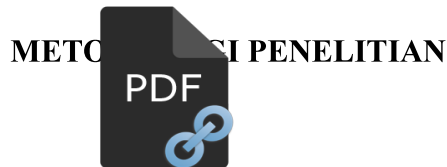
Kerangka Berpikir berfungsi sebagai landasan untuk melakukan pendekatan penelitian yang bersumber dari fakta, pengamatan dan tinjauan Pustaka [26]. Memiliki kerangka berpikir yang jelas memungkinkan penelitian diatur dengan mudah. Adapun kerangka berpikir dapat dilihat pada gambar 2.1 di bawah ini:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



Gambar 2.1. Kerangka Berpikir



3.1 Metode Penelitian

Penelitian ini menggunakan metode kuantitatif, metode kuantitatif mengacu pada nilai numerik atau nominal dan sering digunakan pada penelitian survei atau polling [27]. Penelitian ini berfokus pada analisis data yang dikumpulkan dari tweet pengguna Twitter (X). Data ini akan diolah melalui klasifikasi menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor* dan *Artificial Neural Network*.

3.2 Metode Pengumpulan Data

Pengumpulan data merupakan proses untuk mengumpulkan informasi dari berbagai sumber guna menjawab pertanyaan penelitian atau mendukung dalam pengambilan Keputusan. Berikut merupakan metode pengumpulan data yang digunakan dalam melakukan penelitian, terutama mengenai analisis sentimen menggunakan algoritma *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor* dan *Artificial Neural Network*.

3.2.1 Studi Pustaka

Studi Pustaka merupakan proses pengumpulan data yang dilakukan dengan cara mengumpulkan informasi dari berbagai sumber seperti buku, artikel, jurnal ilmiah, dan sumber lainnya dengan tujuan untuk memperkuat landasan teori dan konteks penelitian yang sedang dilakukan.

3.2.2 Data Primer

Data primer merupakan data yang diambil atau didapatkan langsung dari sumber asli, pertama kali dikumpulkan oleh peneliti. Dalam hal ini, data yang diambil merupakan tweet dari pengguna media sosial Twitter (X) berbahasa Indonesia dengan rentang waktu data disetiap kata kunci.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

3.3 Metode Analisa

Penelitian ini menggunakan metode analisis dengan menerapkan algoritma klasifikasi *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor* dan *Artificial Neural Network* untuk mengevaluasi performa *accuracy*, *precision*, *recall*. Dengan tujuan untuk mengidentifikasi algoritma yang memberikan akurasi terbaik dalam mengklasifikasikan sentimen opini publik mengenai Pilkada Jakarta 2024.

3.4 Tempat dan Waktu Penelitian

3.4.1 Tempat

Penelitian ini dilakukan melalui media sosial Twitter (X) dengan cara crawling tweet pengguna.

3.4.2 Waktu

Waktu penelitian ini dimulai dari bulan september 2024. Waktu kegiatan penelitian disajikan dalam bentuk tabel 3.1 berikut ini:

Tabel 3. 1. Waktu Penelitian

Jenis Kegiatan	Waktu Kegiatan																				
	September				Oktober				November				Desember				Januari				
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
Pengajuan Judul	■																				
Pengumpulan Data		■	■	■	■	■	■	■	■												
Bimbingan Proposal								■	■	■	■	■	■								
Ujian Proposal														■	■						
Pengolahan & Pengujian																■	■	■	■	■	■
Bimbingan Skripsi																					
Ujian Akhir																					■

3.5 Alat dan Bahan

3.5.1 Alat

Dalam melakukan penelitian, diperlukan alat-alat yang menunjang proses penelitian. Adapun alat-alat yang digunakan sebagai berikut :

- a. *Laptop*
- b. *Printer*
- c. Kertas A4 80 gsm

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

d. Tinta Printer

3.5.2 Bahan

Adapun bahan yang digunakan dalam proses penelitian ini sebagai berikut :



- a. Jurnal
- b. Microsoft Office
- c. Google Chrome
- d. Google Colab
- e. Mendeley

3.6 Metode Pengujian dan Pengolahan Data

3.6.1 Metode Pengujian

Metode pengujian dilakukan menggunakan *K-Fold Cross Validation*, *Confusion Matrix* dan ROC AUC sebagai evaluasi algoritma NB, SVM, KNN dan ANN untuk mempresentasikan hasil dari proses klasifikasi.

1. *K-Fold Cross Validation*

K-Fold Cross Validation merupakan jenis uji validasi silang yang dirancang untuk mengevaluasi kinerja proses metode algoritma dengan membagi sampel data secara acak dan mengelompokkan data berdasarkan nilai pemisahan. Kemudian salah satu kelompok k-fold digunakan sebagai data uji dan kelompok sisanya digunakan sebagai data latih [28].

2. *Confusion Matrix*

Confusion matrix adalah Teknik yang digunakan untuk mengukur akurasi dalam klasifikasi. Metode ini dapat mengetahui *accuracy*, *precision* dan *recall* [29]. Berikut adalah model evaluasi menggunakan confusion matrix [30][31]:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Tabel 3.2. *Confusion Matrix*

<i>Actual Class</i>	<i>Predicted Class</i>	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Negative</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

TP: *True Positive* menunjukkan sistem memprediksi positif dan hasilnya benar.

FP: *False Positive* menunjukkan sistem memprediksi positif dan hasilnya salah.

FN: *False Negative* menunjukkan sistem memprediksi negatif dan hasilnya salah.

TN: *True Negative* menunjukkan sistem memprediksi negatif dan hasilnya benar.

Berdasarkan tabel *confusion matrix*, diperoleh nilai *accuracy*, *precision*, dan *recall* sebagai berikut:

1) *Accuracy* menunjukkan tingkat keakuratan klasifikasi model [19]. Rumus *Accuracy* [32]:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

2) *Precision* merupakan kecepatan sistem memberikan hasil dalam menanggapi permintaan informasi pengguna [19]. Rumus *Precision* [32]:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3) *Recall* dapat didefinisikan sebagai ukuran jumlah dokumen teks penting [19]. Rumus *Recall* [32]:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

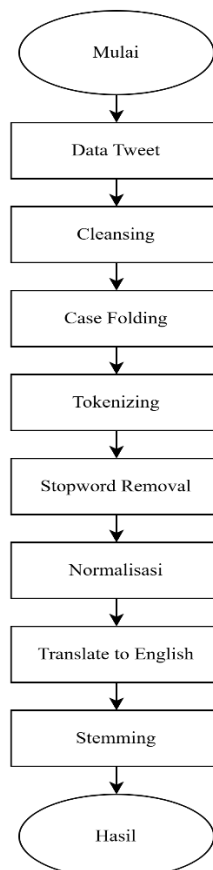
Protected by PDF Anti-Copy Free
 (Upgrade to Pro Version to Remove the Watermark)

3. ROC AUC

ROC (Receiver Operating Characteristic) digunakan untuk menunjukkan data kinerja algoritma klasifikasi dalam bentuk grafik. Kurva ROC dibuat berdasarkan nilai yang telah diperoleh dari perhitungan menggunakan *confusion matrix*, yaitu antara *false positive* dan *true positive*. Untuk membandingkan nilai kinerja setiap algoritma dapat dilakukan dengan mengukur luas di bawah kurva atau AUC (*Area Under Curve*) [33].

3.6.2 Metode Pengolahan Data

Setelah pengumpulan data tweet, selanjutnya dilakukan preprocessing. Tahapan dari *preprocessing* dapat dilihat pada gambar 3.1 berikut:



Gambar 3.1. Tahapan *Preprocessing*

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

1. *Cleansing*

Merupakan tahapan pembersihan data dengan menghapus tanda baca dan non-teks [14]. Seperti url, html, emoji, penomoran dan simbol dan nama pengguna (*username*) [34]. Adapun contoh dari proses cleansing dapat di lihat pada tabel 3.3 berikut:

Tabel 3.3. Contoh Hasil *Cleansing*

Teks	Hasil <i>Cleansing</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA... #MatinyaNKRI https://t.co/FGgnPYTTkR	GUE KAGAK PERDULI AMA PILKADA JAKARTA GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA SIA AKIBAT ULAH ANJING MACAM MEREKA MatinyaNKRI

2. *Case Folding*

Merupakan tahapan mengubah huruf besar menjadi huruf kecil dengan menghilangkan semua tanda baca dalam kalimat [3]. Dengan tujuan untuk memudahkan konsistensi dalam analisis teks [30]. Contoh dari proses *case folding* dapat di lihat pada tabel 3.4 berikut:

Tabel 3.4. Contoh Hasil *Case Folding*

Teks	Hasil <i>Case Folding</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA...	gue kagak perduli ama pilkada jakarta gue gak rela darah kakek dan babe gue pemegang bintang gerilya sia akibat ulah anjing macam mereka matinyankri

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

#MatinyaNKRI

<https://t.co/FGgnPYTTkR>



3. *Tokenizing*

Tokenizing merupakan proses mempersingkat setiap kata dalam dokumen, mengubah karakter yang awalnya huruf besar menjadi huruf kecil. Proses ini memungkinkan teks dibagi menjadi beberapa potongan kecil yang dapat diproses lebih efisien melalui analisis teks [34]. Contoh dari proses *tokenizing* dapat dilihat pada tabel 3.5 berikut:

Tabel 3.5. Contoh Hasil *Tokenizing*

Teks	Hasil <i>Tokenizing</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA... #MatinyaNKRI https://t.co/FGgnPYTTkR	['gue', 'kagak', 'perduli', 'ama', 'pilkada', 'jakarta', 'gue', 'gak', 'rela', 'darah', 'kakek', 'dan', 'babe', 'gue', 'pemegang', 'bintang', 'gerilya', 'sia', 'akibat', 'ulah', 'anjing', 'macam', 'mereka', 'matinyankri']

4. *Stopword Removal*

Pada tahap ini, dilakukan penghapusan kata-kata yang dianggap tidak penting pada saat proses klasifikasi. Hal ini dilakukan dengan menghilangkan kata yang tercantum dalam stoplist seperti kata dan, atau, dengan, di, dari, yang, dll [2]. Contoh dari proses *stopword removal* dapat di lihat pada tabel 3.6 berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Tabel 3. 6. Contoh Hasil *Stopword Removal*

	Hasil <i>Stopword Removal</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA... #MatinyaNKRI	['saya', 'kagak', 'perduli', 'ama', 'pilkada', 'jakarta', 'saya', 'tidak', 'rela', 'darah', 'kakek', 'babe', 'saya', 'pemegang', 'bintang', 'gerilya', 'sia', 'akibat', 'ulah', 'anjing', 'matinyankri']
https://t.co/FGgnPYTTkR	

5. Normalisasi

Merupakan proses pengubahan kata slang yang tidak mengikuti kaidah agar sesuai dengan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI) [14]. Contoh hasil proses normalisasi dapat dilihat pada tabel 3.7 berikut:

Tabel 3. 7. Contoh Hasil Normalisasi

Teks	Hasil Normalisasi
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA... #MatinyaNKRI	['saya', 'kagak', 'perduli', 'ama', 'pilkada', 'jakarta', 'saya', 'tidak', 'rela', 'darah', 'kakek', 'babe', 'saya', 'pemegang', 'bintang', 'gerilya', 'sia', 'akibat', 'ulah', 'anjing', 'matinyankri']
https://t.co/FGgnPYTTkR	

6. *Translate to English*

Proses pengubahan dari Bahasa Indonesia ke Bahasa Inggris agar mempermudah proses labelling dan mendapatkan hasil akurasi yang maksimal. Contoh hasil proses *Translate to English* dapat dilihat pada tabel 3.8 berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Tabel 3. 8. Contoh Hasil *Translate to English*

	Hasil <i>Translate to English</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA...	['aku', 'kagak', 'care', 'ama', 'pilkada', 'jakarta', 'me', 'no', 'willing', 'blood', 'grandfather', 'babe', 'me', 'holder', 'star', 'guerilla', 'vain', 'consequence', 'act', 'dog', 'matinyankri']
#MatinyaNKRI https://t.co/FGgnPYTTkR	

7. *Stemming*

Pada tahap ini imbuhan dihilangkan atau bentuk kata diubah kembali menjadi baku [2]. *Stemming* salah satu proses yang penting dilakukan karena memiliki pengaruh terhadap kualitas hasil analisis [35]. Contoh hasil proses *stemming* dapat dilihat pada tabel 3.9 berikut:

Tabel 3. 9. Contoh Hasil *Stemming*

Teks	Hasil <i>Stemming</i>
GUE KAGAK PERDULI AMA PILKADA JAKARTA.. GUE GAK RELA DARAH KAKEK DAN BABE GUE PEMEGANG BINTANG GERILYA.. SIA2 AKIBAT ULAH ANJING2 MACAM MEREKA...	['aku ', 'kagak ', 'care ', 'ama ', 'pilkada ', 'jakarta ' , 'me ', 'no ', 'will ', 'blood , 'grandfath ', 'babe ', 'me , 'holder ', 'star ', 'guerilla , 'vain ', 'consequ ', 'act ' , 'dog ', 'matinyankri ']
#MatinyaNKRI https://t.co/FGgnPYTTkR	

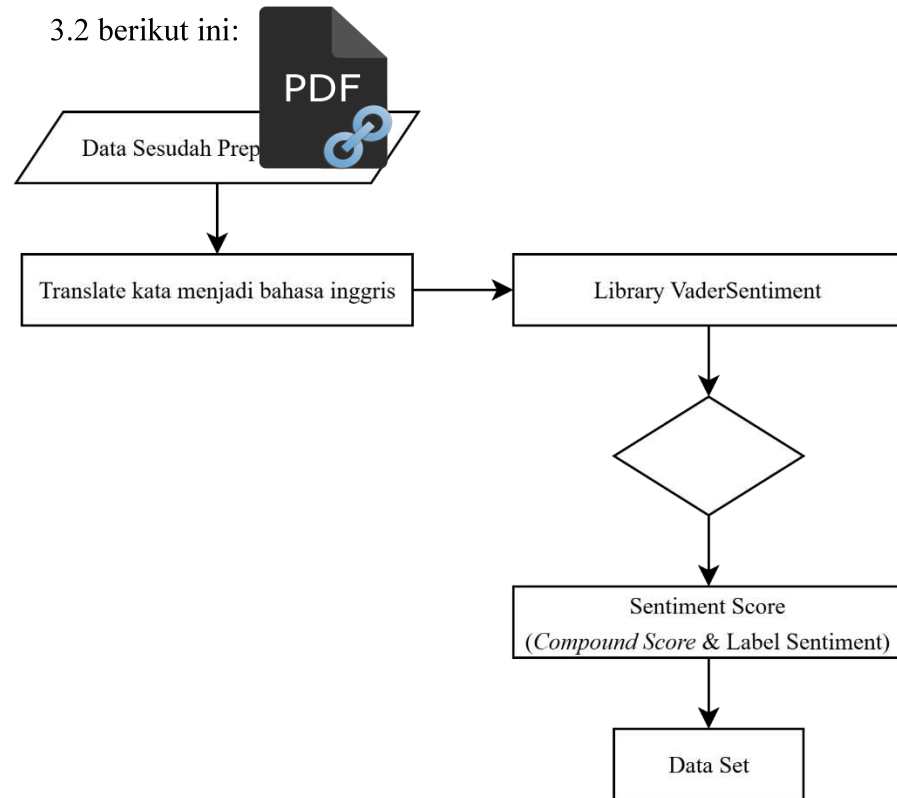
Setelah dilakukan *preprocessing*, dilanjutkan dengan proses labeling menggunakan Vader. Vader peka terhadap polaritas dan kekuatan emosi, mengenali sentimen positif maupun negatif, dan dirancang untuk konten media sosial yang biasanya menampilkan bahasa tidak baku, termasuk tanda baca terbatas, akronim, emotikon,

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

dan bahasa gaul. Adapun tahapan labeling dapat dilihat pada gambar

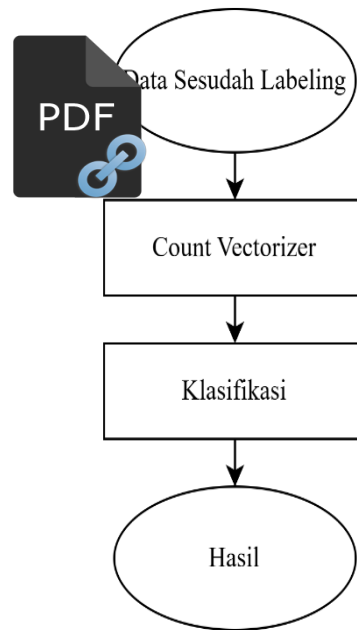
3.2 berikut ini:



Gambar 3.2. Tahapan Labeling

Sebelum menggunakan kumpulan data sentimen, untuk pembuatan model, penting untuk mengubah kumpulan data tersebut ke dalam format matriks numerik menggunakan *Count Vectorizer*. Selanjutnya dapat dilakukan klasifikasi model analisis sentimen menggunakan *Naïve Bayes*, *Support Vector Machine*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*. Proses labeling ini menggunakan *VaderSentiment*, Adapun alur dari proses *CountVectorizer* sebagai berikut:

Protected by PDF Anti-Copy Free
(Upgrade to Pro Version to Remove the Watermark)



Gambar 3.3. Tahapan *CountVectorizer* dan Klasifikasi

HASIL PENELITIAN DAN PEMBAHASAN



4.1 Gambaran Umum

Twitter merupakan salah satu platform media sosial yang paling banyak digunakan di kalangan pengguna internet, berkat desainnya yang lugas dan antarmuka yang ramah pengguna, yang memungkinkan setiap orang membagikan pemikiran mereka secara bebas [36]. Pada platform tersebut, pengguna dapat menggunakan 280 karakter untuk berinteraksi dengan pengguna lainnya, termasuk memberikan opini terhadap isu-isu yang sedang ramai diperbincangkan.

Tweet sering kali mengungkapkan pendapat yang jelas, yang membantu algoritma lebih mudah mengidentifikasi sentimen teks sebagai positif, negatif, atau netral. Analisis sentimen terhadap opini publik mengenai Pilkada Jakarta 2024 di media sosial berperan penting untuk memahami dinamika sosial dan politik dalam Masyarakat. Hasil analisis yang akurat memungkinkan semua pemangku kebijakan, baik Lembaga pemerintah, kandidat, maupun Masyarakat untuk membuat Keputusan yang lebih tepat yang menjunjung tinggi integritas demokrasi dan mempromosikan keharmonisan sosial.

4.2 Hasil

4.2.1 Pengumpulan Data

Pengumpulan data dilakukan dengan cara crawling tweet menggunakan tool *tweet-harvest* yang memerlukan twitter auth_token sebagai kode akses. Pengumpulan data dilakukan berdasarkan rentang waktu Januari 2024-Januari 2025 dengan keyword “Pilkada Jakarta 2024” sebanyak 26.354 dikeyword, Adapun hasil dari crawling sebagai berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name
1.867220e+18	Thu Dec 12 14:50:46 +0000 2024		Ridwan Kamil Total di Pilkada	1.867220e+18	NaN	NaN
1.867200e+18	Thu Dec 12 13:22:42 +0000 2024			1.867200e+18	https://pbs.twimg.com/ext_tw_video_thumb/18671...	NaN
1.865610e+18	Sun Dec 08 11:51:24 +0000 2024	7.0	Bujer-bujer Hasto sekarang pada orgas...	1.865730e+18	NaN	aewin86
1.863880e+18	Tue Dec 03 09:21:08 +0000 2024	20.0	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	1.863880e+18	https://pbs.twimg.com/media/Gd3JWXca0AAM38N.jpg	NaN

Gambar 4.1. Hasil Crawling

4.2.2 Preprocessing

Sumber data yang berasal dari media sosial, sering kali tidak teratur dan dipenuhi banyak *noise* yang bisa memengaruhi hasil analisis. Maka dari itu, diperlukan preprocessing terlebih dahulu dengan cara pembersihan atau pengolahan pada data mentah agar dapat dilakukan pelabelan, pemodelan dan evaluasi.

4.2.2.1 Cleansing

Proses cleansing merupakan tahapan pembersihan data dari karakter yang tidak diperlukan dalam klasifikasi dengan cara menghapus tanda baca dan karakter non-teks. Seperti url, html, emoji, penomoran, simbol-simbol dan nama pengguna (*username*). Adapun hasil cleansing sebagai berikut:

	full_text	cleansing
0	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...
2	@aewin86 Bujer-bujer Hasto sekarang pada orgas...	Bujerbujer Hasto sekarang pada orgasme eufori...
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...
4	Potret nyata kegagalan parpol dalam mengemban ...	Potret nyata kegagalan parpol dalam mengemban ...
5	Dari ramainya seteru pilkada Jateng & Jaka...	Dari ramainya seteru pilkada Jateng amp Jakart...
6	Kata kader PKS Suswono: Singapura Ketar-Ketir ...	Kata kader PKS Suswono Singapura KetarKetir ji...
7	Kekanak-Kanakan Pihak yang mengglorifikasi keg...	KekanakKanakan Pihak yang mengglorifikasi kega...
8	Bukti Kegagalan RK Dalam Rancangan Halte Di Ba...	Bukti Kegagalan RK Dalam Rancangan Halte Di Ba...
9	@FelixSGL1 Yap. Tapi yang lumayan mengejutkan ...	Yap Tapi yang lumayan mengejutkan adalah suar...

Gambar 4.2. Hasil Cleansing

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

4.2.2.2 Case Folding

Case folding merupakan tahapan pengubahan huruf besar menjadi huruf kecil. Dengan tujuan untuk memudahkan konsistensi dalam analisis teks. Adapun hasil *case folding* dapat dilihat sebagai berikut:

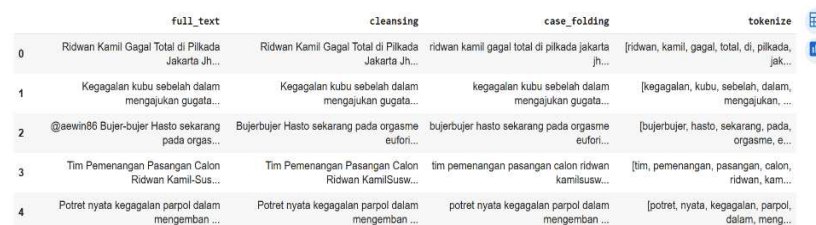


	full_text	cleansing	case_folding
0	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	ridwan kamil gagal total di pilkada jakarta jh...
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...
2	@aewin86 Bujer-bujer Hasto sekarang pada orgas...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...
4	Potret nyata kegagalan parpol dalam mengemban ...	Potret nyata kegagalan parpol dalam mengemban ...	potret nyata kegagalan parpol dalam mengemban ...

Gambar 4. 3. Hasil *Case Folding*

4.2.2.3 Tokenizing

Tokenizing adalah tahapan membagi teks menjadi potongan kecil yang dapat diproses lebih efisien melalui analisis teks. Adapun hasil dari tokenizing dapat dilihat sebagai berikut:



	full_text	cleansing	case_folding	tokenize
0	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	ridwan kamil gagal total di pilkada jakarta jh...	[ridwan, kamil, gagal, total, di, pilkada, jak...
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	[kegagalan, kubu, sebelah, dalam, mengajukan, ...
2	@aewin86 Bujer-bujer Hasto sekarang pada orgas...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	[bujerbujer, hasto, sekarang, pada, orgasme, e...
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...	[tim, pemenangan, pasangan, calon, ridwan, kam...
4	Potret nyata kegagalan parpol dalam mengemban ...	Potret nyata kegagalan parpol dalam mengemban ...	potret nyata kegagalan parpol dalam mengemban ...	[potret, nyata, kegagalan, parpol, dalam, meng...

Gambar 4.4. Hasil *Tokenizing*

4.2.2.4 Stopword Removal

Tahapan ini dilakukan dengan menghilangkan kata yang tercantum dalam stoplist seperti kata dan, atau, dengan, di, dari, yang, dll. Adapun hasil dari stopword removal dapat dilihat sebagai berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

	full_text	cleansing	case_folding	tokenize	stopword_removal
0	Ridwan Kamil Gagah Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagah Total di Pilkada Jakarta Jh...	ridwan kamil gagah total di pilkada jakarta jh...	[ridwan, kamil, gagah, total, di, pilkada, jak...	[ridwan, kamil, gagah, total, pilkada, jakarta...
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	[kegagalan, kubu, sebelah, dalam, mengajukan, ...	[kegagalan, kubu, sebelah, mengajukan, gugatan...
2	@aewin86 Bujer-bujer Hasto sekarang pada orgasme eufori...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	[bujerbujer, hasto, sekarang, pada, orgasme, e...	[bujerbujer, hasto, orgasme, euforia, pramono...
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...	[tim, pemenangan, pasangan, calon, ridwan, kam...	[tim, pemenangan, pasangan, calon, ridwan, kam...
4	Potret nyata kegagalan parpol dalam mengemban ...	Potret nyata kegagalan parpol dalam mengemban ...	potret nyata kegagalan parpol dalam mengemban ...	[potret, nyata, kegagalan, parpol, dalam, meng...	[potret, nyata, kegagalan, parpol, mengemban, ...

Gambar 4.5. Hasil *Stopword Removal*

4.2.2.5 Normalisasi

Merupakan proses perubahan kata slang yang tidak mengikuti kaidah agar sesuai dengan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI). Adapun hasil dari normalisasi sebagai berikut:

	full_text	cleansing	case_folding	tokenize	stopword_removal	normalisasi
0	Ridwan Kamil Gagah Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagah Total di Pilkada Jakarta Jh...	ridwan kamil gagah total di pilkada jakarta jh...	[ridwan, kamil, gagah, total, di, pilkada, jak...	[ridwan, kamil, gagah, total, pilkada, jakarta...	[ridwan, kamil, gagah, total, pilkada, jakarta...
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	[kegagalan, kubu, sebelah, dalam, mengajukan, ...	[kegagalan, kubu, sebelah, mengajukan, gugatan...	[kegagalan, kubu, sebelah, mengajukan, gugatan...
2	@aewin86 Bujer-bujer Hasto sekarang pada orgasme eufori...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	[bujerbujer, hasto, sekarang, pada, orgasme, e...	[bujerbujer, hasto, orgasme, euforia, pramono...	[bujerbujer, hasto, orgasme, euforia, pramono...
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...	[tim, pemenangan, pasangan, calon, ridwan, kam...	[tim, pemenangan, pasangan, calon, ridwan, kam...	[tim, pemenangan, pasangan, calon, ridwan, kam...
4	Potret nyata kegagalan parpol dalam mengemban ...	Potret nyata kegagalan parpol dalam mengemban ...	potret nyata kegagalan parpol dalam mengemban ...	[potret, nyata, kegagalan, parpol, dalam, meng...	[potret, nyata, kegagalan, parpol, mengemban, ...	[potret, nyata, kegagalan, parpol, mengemban, ...

Gambar 4.6. Hasil Normalisasi

4.2.2.6 Translate To English

Translate to English merupakan proses perubahan kata menggunakan Bahasa Indonesia diubah menjadi bahasa Inggris agar dapat dilakukan pelabelan menggunakan Vader. VADER merupakan model untuk menganalisis sentimen yang didasarkan pada aturan yang memanfaatkan kamus kosakata dan ungkapan dalam Bahasa Inggris. Sehingga model ini tidak bisa secara langsung mengerti atau menangani teks dalam Bahasa Indonesia tanpa adanya terjemahan. Adapun hasil dari translate to English dapat dilihat sebagai berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

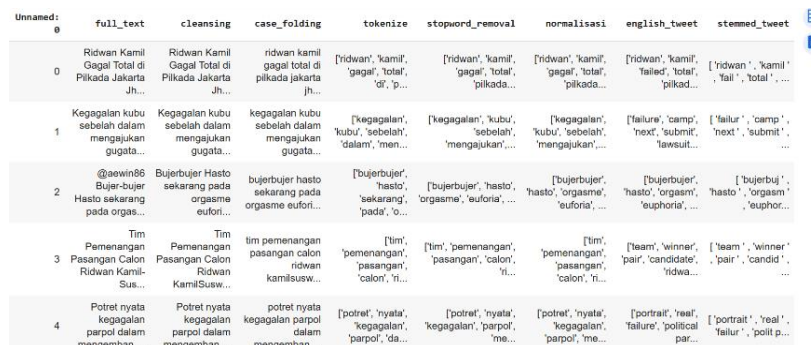


Unnamed: 0	full_text	cleansing	case_folding	tokenize	stopword_removal	normalisasi	english_tweet
0	Ridwan Kamil Gagagal Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagagal Total di Pilkada Jakarta Jh...	ridwan kamil gagagal total di pilkada jakarta jh...	['ridwan', 'kamil', 'gagagal', 'total', 'di', 'pilkada', 'j', 'h...']	['ridwan', 'kamil', 'gagagal', 'total', 'pilkada...']	['ridwan', 'kamil', 'gagagal', 'total', 'pilkada...']	['ridwan', 'kamil', 'failed', 'total', 'pilkada...']
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	['kegagalan', 'kubu', 'sebelah', 'dalam', 'men...']	['kegagalan', 'kubu', 'sebelah', 'mengajukan...']	['kegagalan', 'kubu', 'sebelah', 'mengajukan...']	['failure', 'camp', 'next', 'submit', 'lawsuit...']
2	@sewin86 Bujerbujer Hasto sekarang pada orgasme eufori...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	['bujerbujer', 'hasto', 'sekarang', 'pada', 'o...']	['bujerbujer', 'hasto', 'orgasme', 'euforia', ...]	['bujerbujer', 'hasto', 'orgasme', 'euforia', ...]	['bujerbujer', 'hasto', 'orgasm', 'euphoria', ...]
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['team', 'winner', 'pair', 'candidate', 'ridwa...']
4	Potret nyata kegagalan parpol dalam mengemban...	Potret nyata kegagalan parpol dalam mengemban...	potret nyata kegagalan parpol dalam mengemban...	['potret', 'nyata', 'kegagalan', 'parpol', 'da...']	['potret', 'nyata', 'kegagalan', 'parpol', 'me...']	['potret', 'nyata', 'kegagalan', 'parpol', 'me...']	['portrait', 'real', 'failure', 'political par...']

Gambar 4.7. Hasil *Translate to English*

4.2.2.7 Stemming

Pada tahap ini imbuhan dihilangkan atau bentuk kata diubah Kembali menjadi baku. *Stemming* salah satu proses yang penting dilakukan karena memiliki pengaruh terhadap kualitas hasil analisis. Adapun hasil stemming dapat dilihat sebagai berikut:



Unnamed: 0	full_text	cleansing	case_folding	tokenize	stopword_removal	normalisasi	english_tweet	stemmed_tweet
0	Ridwan Kamil Gagagal Total di Pilkada Jakarta Jh...	Ridwan Kamil Gagagal Total di Pilkada Jakarta Jh...	ridwan kamil gagagal total di pilkada jakarta jh...	['ridwan', 'kamil', 'gagagal', 'total', 'di', 'pilkada', 'j', 'h...']	['ridwan', 'kamil', 'gagagal', 'total', 'pilkada...']	['ridwan', 'kamil', 'gagagal', 'total', 'pilkada...']	['ridwan', 'kamil', 'failed', 'total', 'pilkada...']	['ridwan', 'kamil', 'fail', 'total', ...]
1	Kegagalan kubu sebelah dalam mengajukan gugata...	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	['kegagalan', 'kubu', 'sebelah', 'dalam', 'men...']	['kegagalan', 'kubu', 'sebelah', 'mengajukan...']	['kegagalan', 'kubu', 'sebelah', 'mengajukan...']	['failure', 'camp', 'next', 'submit', 'lawsuit...']	['failur', 'camp', 'next', 'submit', 'lawsuit...']
2	@sewin86 Bujerbujer Hasto sekarang pada orgasme eufori...	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	['bujerbujer', 'hasto', 'sekarang', 'pada', 'o...']	['bujerbujer', 'hasto', 'orgasme', 'euforia', ...]	['bujerbujer', 'hasto', 'orgasme', 'euforia', ...]	['bujerbujer', 'hasto', 'orgasme', 'euphoria', ...]	['bujerbuj', 'hasto', 'orgasme', 'euphor...']
3	Tim Pemenangan Pasangan Calon Ridwan Kamil-Sus...	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilsusw...	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...']	['team', 'winner', 'pair', 'candidate', 'ridwa...']	['team', 'winner', 'pair', 'candid', ...]
4	Potret nyata kegagalan parpol dalam mengemban...	Potret nyata kegagalan parpol dalam mengemban...	potret nyata kegagalan parpol dalam mengemban...	['potret', 'nyata', 'kegagalan', 'parpol', 'da...']	['potret', 'nyata', 'kegagalan', 'parpol', 'me...']	['potret', 'nyata', 'kegagalan', 'parpol', 'me...']	['portrait', 'real', 'failure', 'political par...']	['portrait', 'real', 'failur', 'polit p...']

Gambar 4.8. Hasil *Stemming*

4.2.3 Labelling

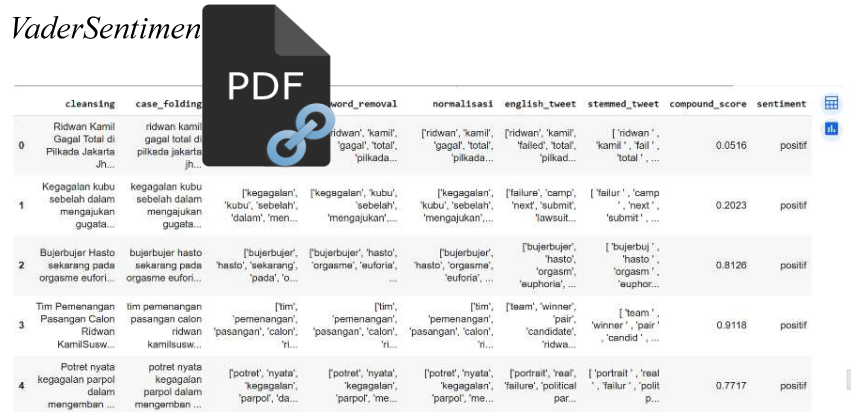
Labeling merupakan tahapan pemberian label sentiment apakah positif, negatif dan netral. Proses labeling menggunakan library *VaderSentiment*. Penggunaan *VaderSentiment* harus menggunakan teks Bahasa Inggris. Proses labeling menghasilkan nilai *compound score* dengan nilai <0 memiliki nilai sentiment negative, $=0$ memiliki nilai sentiment netral, dan >0 memiliki nilai

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

sentiment positif. Berikut hasil labeling menggunakan

VaderSentimen



	cleansing	case_folding	word_removal	normalisasi	english_tweet	stemmed_tweet	compound_score	sentiment
0	Ridwan Kamil Gagal Total di Pilkada Jakarta Jh...	ridwan kamil gagal total di pilkada jakarta jh...	[ridwan', 'kamil', 'gagal', 'total', 'pilkada', ...	['ridwan', 'kamil', 'gagal', 'total', 'pilkada', ...	['ridwan', 'kamil', 'failed', 'total', 'pilkad', ...	['ridwan', 'fail', 'kamil', 'total', ...	0.0516	positif
1	Kegagalan kubu sebelah dalam mengajukan gugata...	kegagalan kubu sebelah dalam mengajukan gugata...	['kegagalan', 'kubu', 'sebelah', 'dalam', 'men...', 'mengajukan', ...	['kegagalan', 'kubu', 'sebelah', 'mengajukan', ...	['failure', 'camp', 'next', 'submit', 'lawsuit', ...	['failur', 'camp', 'next', 'submit', ...	0.2023	positif
2	Bujerbujer Hasto sekarang pada orgasme eufori...	bujerbujer hasto sekarang pada orgasme eufori...	['bujerbujer', 'hasto', 'sekarang', 'pada', 'o...', ...	['bujerbujer', 'hasto', 'orgasme', 'euforia', '...', ...	['bujerbujer', 'hasto', 'orgasme', 'euforia', ...	['bujerbu', 'hasto', 'orgasm', 'euphor', ...	0.8126	positif
3	Tim Pemenangan Pasangan Calon Ridwan KamilSusw...	tim pemenangan pasangan calon ridwan kamilSusw...	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...', ...	['tim', 'pemenangan', 'pasangan', 'calon', 'ri...', ...	['team', 'winner', 'pair', 'candidate', 'ridwa...', ...	['team', 'winner', 'pair', 'candid', ...	0.9118	positif
4	Potret nyata kegagalan parpol dalam mengamban ...	potret nyata kegagalan parpol dalam mengamban ...	['potret', 'nyata', 'kegagalan', 'parpol', 'da...', ...	['potret', 'nyata', 'kegagalan', 'parpol', 'me...', ...	['portrait', 'real', 'failure', 'political', 'par...', ...	['portrait', 'real', 'failure', 'polit', 'p...', ...	0.7717	positif

Gambar 4.9. Hasil Labelling

Adapun komposisi sentimen dari proses labeling sebagai berikut:

```

jumlah sentimen :
sentiment
positif      3367
netral       1230
negatif      1044
Name: count, dtype: int64

```

Gambar 4. 10. Hasil Komposisi Sentimen

4.2.4 Visualisasi

Visualisasi merupakan langkah untuk menampilkan data atau informasi dalam format yang visual atau grafis, seperti chart, diagram, atau peta, agar lebih mudah dipahami, dianalisis, dan disampaikan. Visualisasi data menggunakan pie chart dan wordcloud. Hasil representasi sentimen pie chart sebagai berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

4.2.5 *CountVectorizer*

CountVectorizer merupakan metode untuk menentukan frekuensi setiap kata dalam suatu dokumen, menghasilkan representasi vektor yang memperlihatkan angka kemunculan masing-masing kata. Menghitung frekuensi kata adalah teknik yang sederhana, namun jika beberapa kata muncul berulang kali di banyak dokumen, angka frekuensi yang tinggi mungkin tidak memberikan makna yang signifikan dalam vektor relevansinya.

4.2.6 Klasifikasi

Pada penelitian ini, dilakukan klasifikasi dengan menerapkan algoritma *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*. Pemisahan antara data pelatihan dan pengujian memanfaatkan teknik *Cross-Validation k-10*, di mana satu bagian data dipakai sebagai data uji, sementara sembilan bagian lainnya digunakan sebagai data pelatihan. Proses ini diulang sebanyak sepuluh kali untuk memastikan bahwa evaluasi yang dilakukan bersifat konsisten dan mencerminkan keadaan yang sebenarnya.

4.2.7 Evaluasi

Evaluasi hasil klasifikasi terdiri dari *cross validation k-10*, *Confusion Matrix* dan grafik ROC-AUC, memberikan wawasan yang lebih baik mengenai efektivitas model yang diterapkan menggunakan *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*. Adapun hasil evaluasi dari keenam algoritma dapat dilihat sebagai berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

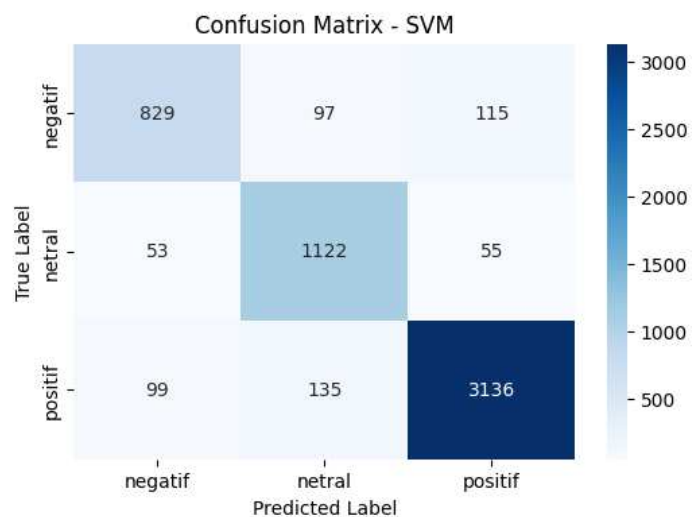
4.2.7.1 Support Vector Machine

Evaluasi hasil cross validation k-10 mendapatkan rata-rata 90%, precision 90%, recall 90%, f1-score 90%, accuracy 90% dan ROC AUC 91%. Adapun hasil dari algoritma *Support Vector Machine* dapat dilihat sebagai berikut:

↳ Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.900885	0.861064	0.879475	0.868762
1	0.890071	0.854717	0.885131	0.866871
2	0.920213	0.906036	0.910473	0.907545
3	0.907801	0.873075	0.885551	0.878784
4	0.891844	0.861490	0.864135	0.862288
5	0.882979	0.867166	0.846368	0.854086
6	0.898936	0.882143	0.857878	0.867436
7	0.909574	0.880088	0.883296	0.880809
8	0.911348	0.873880	0.896398	0.884281
9	0.904255	0.877070	0.887479	0.881357
Rata-rata	0.901791	0.873673	0.879619	0.875222

Gambar 4. 13. Hasil Evaluasi *Cross Validation K-10 SVM*



Gambar 4. 14. Hasil Evaluasi *Confusion Matrix SVM*

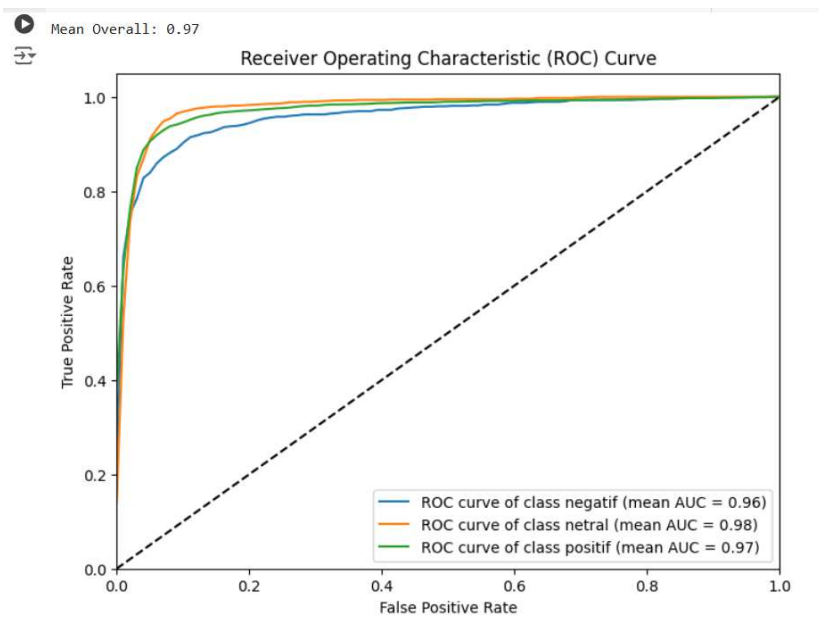
Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.85	0.80	0.82	1041
netral	0.83	0.91	0.87	1230
positif	0.95	0.93	0.94	3370
accuracy			0.90	5641
macro avg	0.87	0.88	0.88	5641
weighted avg	0.90	0.90	0.90	5641

Gambar 4.15. Hasil *Classification Report* SVM



Gambar 4.16. Grafik ROC AUC SVM

4.2.7.2 *Naïve Bayes*

Evaluasi dari cross validation k-10 mendapatkan rata-rata 79%, precision 79%, recall 79%, f1-score 78%, accuracy 79% dan ROC AUC 92%. Adapun hasil dari algoritma *Naïve Bayes* dapat dilihat sebagai berikut:

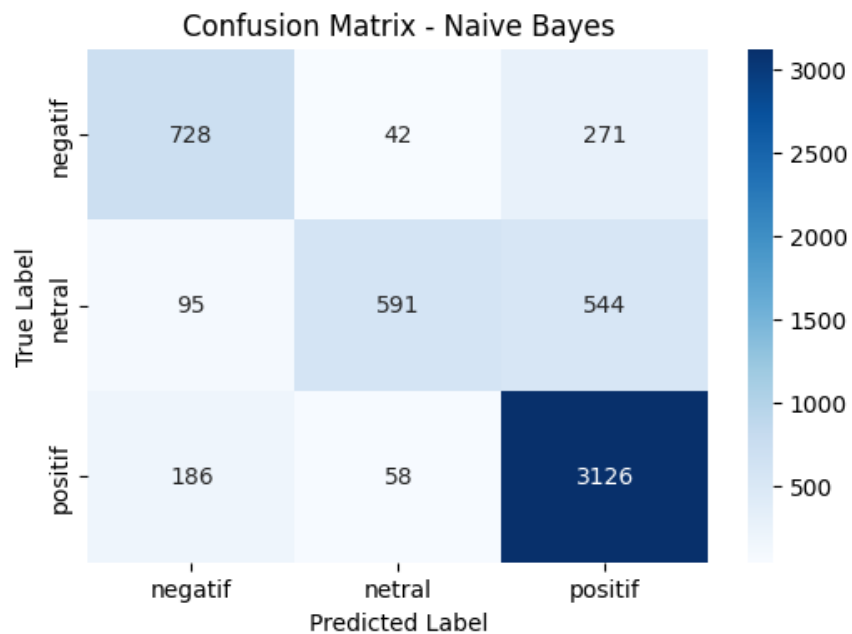
Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.783671	0.783793	0.697321	0.716202
1	0.778376	0.812460	0.745762	0.770246
2	0.783651	0.796238	0.723522	0.745353
3	0.783649	0.824071	0.741201	0.764205
4	0.783688	0.774832	0.691302	0.713703
5	0.778369	0.782234	0.691188	0.717192
6	0.757092	0.773586	0.669195	0.693890
7	0.767730	0.759995	0.676068	0.695215
8	0.797872	0.796262	0.700681	0.722426
9	0.789007	0.791521	0.690668	0.722486
Rata-rata	0.787982	0.789499	0.702691	0.726092

Gambar 4. 17. Hasil Evaluasi *Cross Validation K-10 Naive Bayes*



Gambar 4. 18. Hasil Evaluasi *Confusion Matrix Naive Bayes*

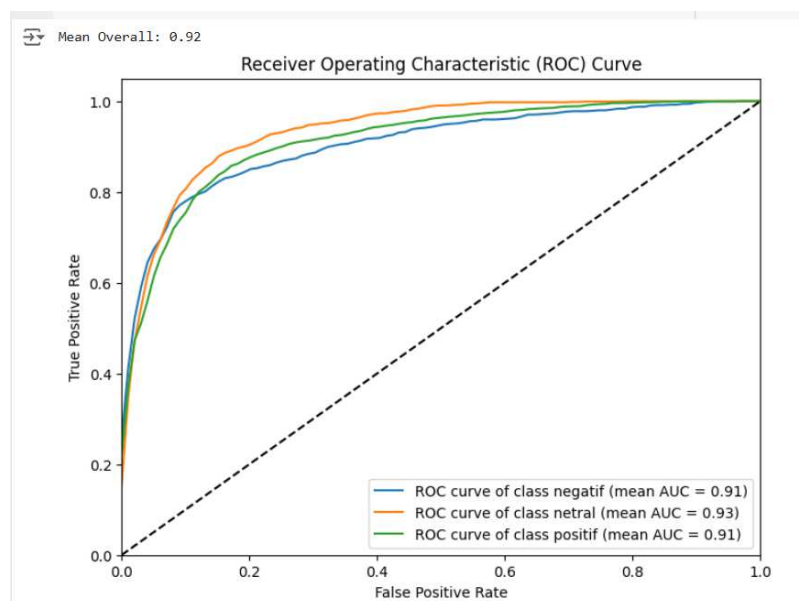
Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.72	0.70	0.71	1041
netral	0.86	0.48	0.62	1230
positif	0.79	0.93	0.86	3370
accuracy			0.79	5641
macro avg	0.79	0.70	0.73	5641
weighted avg	0.79	0.79	0.78	5641

Gambar 4. 19. Hasil *Classification Report Naive Bayes*



Gambar 4. 20. Grafik ROC AUC *Naive Bayes*

4.2.7.3 *K-Nearest Neighbor*

Evaluasi dari cross validation k-10 mendapatkan rata-rata 70%, precision 77%, recall 70%, f1-score 71%, accuracy 70% dan ROC AUC 83%. Adapun hasil dari algoritma *K-Nearest Neighbor* dapat dilihat sebagai berikut:

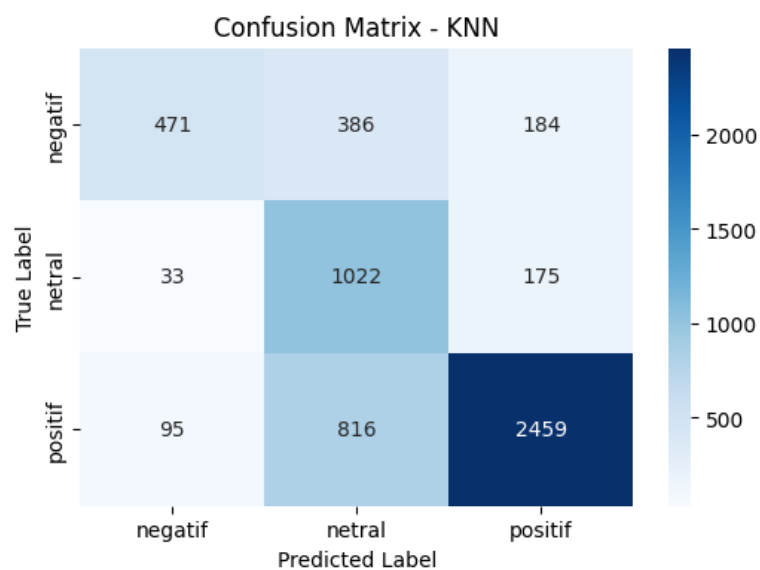
Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.69823	0.718226	0.709353	0.688403
1	0.70489	0.705803	0.693674	0.655133
2	0.70723	0.755769	0.683837	0.682892
3	0.69743	0.698927	0.642888	0.627535
4	0.69305	0.681667	0.649644	0.632768
5	0.670213	0.709898	0.648355	0.625215
6	0.709220	0.713670	0.665519	0.657078
7	0.705674	0.697098	0.678616	0.657009
8	0.691489	0.672796	0.663736	0.634975
9	0.703901	0.718099	0.678552	0.669467
Rata-rata	0.700578	0.707195	0.671417	0.653047

Gambar 4. 21. Hasil Evaluasi *Cross Validation K-10 KNN*



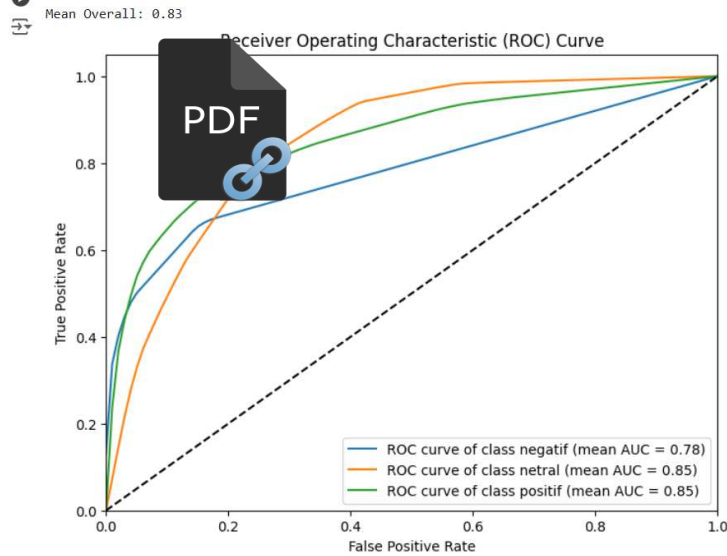
Gambar 4. 22. Hasil Evaluasi *Confusion Matrix KNN*

Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.79	0.45	0.57	1041
netral	0.46	0.83	0.59	1230
positif	0.87	0.73	0.79	3370
accuracy			0.70	5641
macro avg	0.71	0.67	0.65	5641
weighted avg	0.77	0.70	0.71	5641

Gambar 4. 23. Hasil *Classification Report KNN*

Protected by PDF Anti-Copy Free
 (Upgrade to Pro Version to Remove the Watermark)



Gambar 4. 24. Grafik ROC AUC KNN

4.2.7.4 Artificial Neural Network

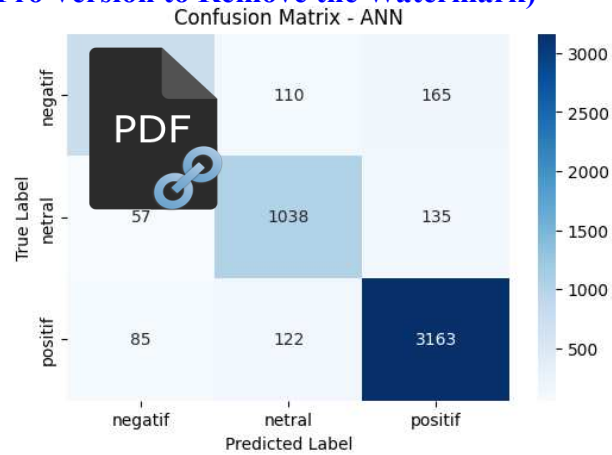
Evaluasi dari cross validation k-10 mendapatkan rata-rata 88%, precision 88%, recall 88%, f1-score 88%, accuracy 88% dan ROC AUC 96%. Adapun hasil dari algoritma *Artificial Neural Network* dapat dilihat sebagai berikut:

Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.874336	0.831626	0.833490	0.832510
1	0.882979	0.854812	0.858183	0.853958
2	0.886525	0.869363	0.859918	0.864077
3	0.884752	0.862023	0.834433	0.847164
4	0.875887	0.852042	0.838192	0.842607
5	0.875887	0.860884	0.825794	0.840331
6	0.868794	0.864272	0.814837	0.835270
7	0.881206	0.857671	0.830942	0.841689
8	0.897163	0.872649	0.862950	0.866454
9	0.877660	0.854212	0.835724	0.843916
Rata-rata	0.880519	0.857955	0.839446	0.846798

Gambar 4. 25. Hasil Evaluasi *Cross Validation K-10* ANN

Protected by PDF Anti-Copy Free
(Upgrade to Pro Version to Remove the Watermark)

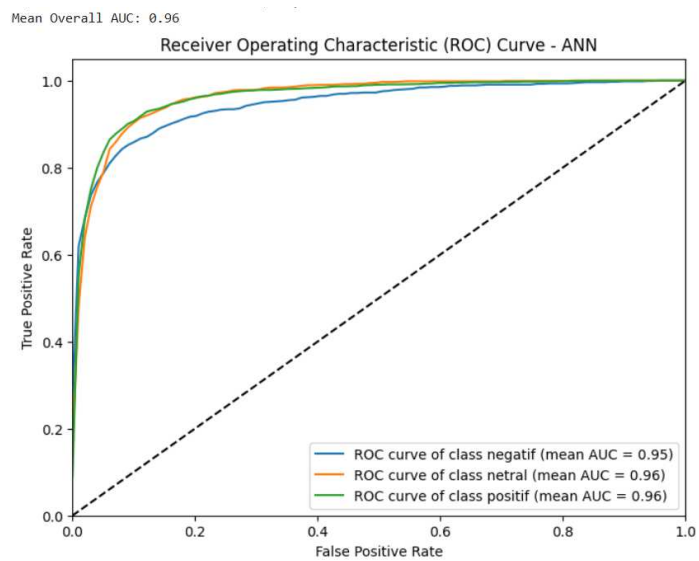


Gambar 4. 26. Hasil Evaluasi *Confusion Matrix* ANN

Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.84	0.74	0.79	1041
netral	0.82	0.84	0.83	1230
positif	0.91	0.94	0.93	3370
accuracy			0.88	5641
macro avg	0.86	0.84	0.85	5641
weighted avg	0.88	0.88	0.88	5641

Gambar 4. 27. Hasil *Classification Report* ANN



Gambar 4. 28. Grafik ROC AUC ANN

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

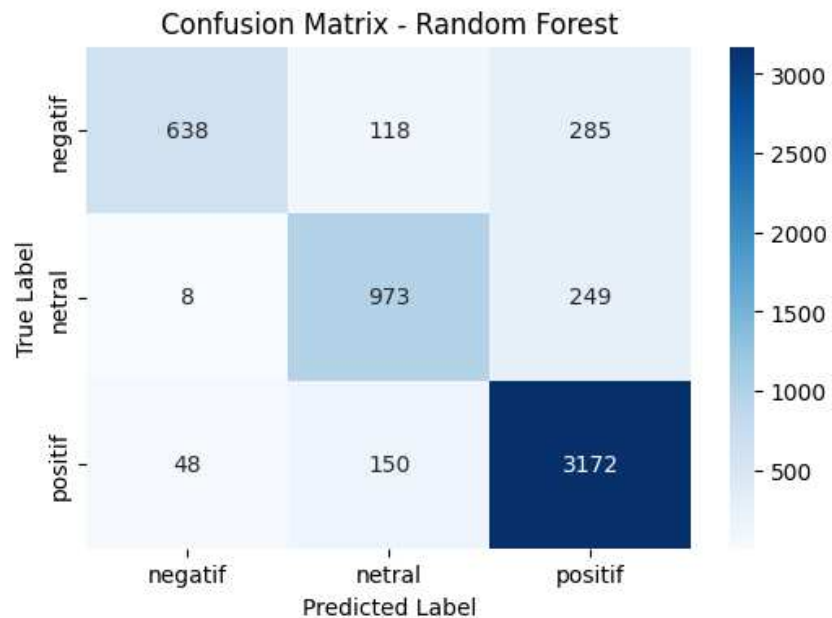
4.2.7.5 Random Forest

Evaluasi hasil validation k-10 mendapatkan rata-rata 84%, precision 84%, recall 85%, f1-score 84%, accuracy 85% dan ROC AUC 90%. Adapun hasil dari algoritma *Random Forest* dapat dilihat sebagai berikut:

Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.861947	0.858325	0.792422	0.817407
1	0.828014	0.815780	0.779017	0.787708
2	0.852837	0.858967	0.788789	0.812570
3	0.847518	0.840061	0.776230	0.800724
4	0.845745	0.849478	0.781648	0.804430
5	0.817376	0.842061	0.739145	0.771532
6	0.851064	0.875708	0.775756	0.806910
7	0.867021	0.875846	0.802733	0.829012
8	0.865248	0.862140	0.804946	0.827025
9	0.842199	0.852524	0.778708	0.802888
Rata-rata	0.847897	0.853089	0.781940	0.806021

Gambar 4. 29. Hasil Evaluasi *Cross Validation K-10 Random Forest*



Gambar 4. 30. Hasil Evaluasi *Confusion Matrix Random Forest*

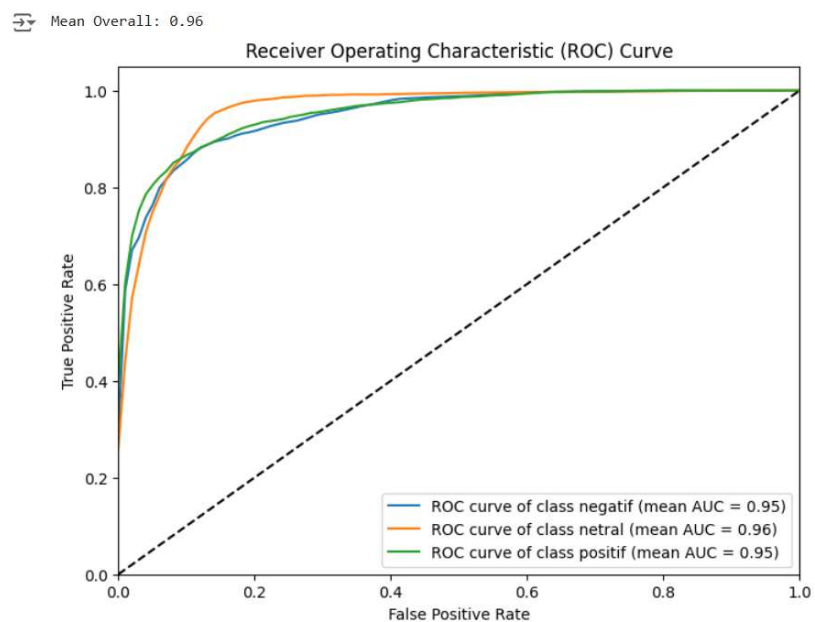
Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.92	0.61	0.74	1041
netral	0.78	0.79	0.79	1230
positif	0.86	0.94	0.90	3370
accuracy			0.85	5641
macro avg	0.85	0.78	0.81	5641
weighted avg	0.85	0.85	0.84	5641

Gambar 4. 31. Hasil *Classification Report Random Forest*



Gambar 4. 32. Grafik ROC AUC *Random Forest*

4.2.7.6 Decision Tree

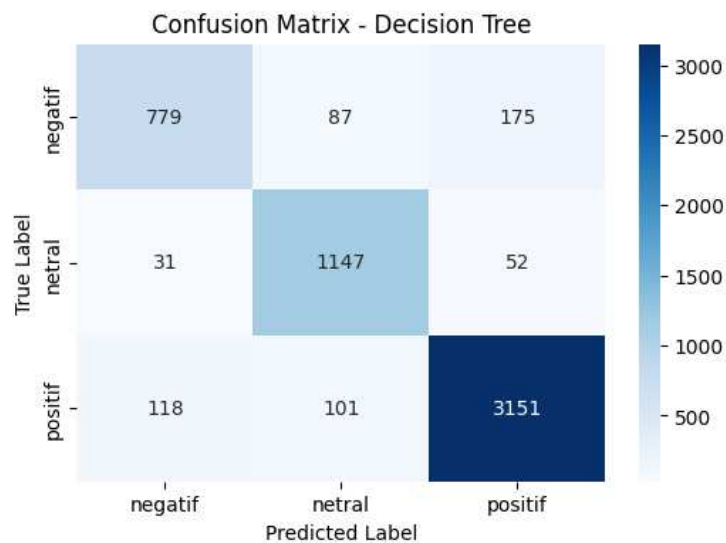
Evaluasi dari cross validation k-10 mendapatkan rata-rata 90%, precision 90%, recall 90%, f1-score 90%, accuracy 90% dan ROC AUC 97%. Adapun hasil dari algoritma *Decision Tree* dapat dilihat sebagai berikut:

Protected by PDF Anti-Copy Free
 (Upgrade to Pro Version to Remove the Watermark)

Hasil Cross Validation:

	Accuracy	Precision	Recall	F1 Score
0	0.8814	0.888219	0.888121	0.888021
1	0.8525	0.857027	0.873288	0.862854
2	0.8482	0.879936	0.881889	0.880526
3	0.8517	0.853923	0.866931	0.859526
4	0.89121	0.899449	0.890366	0.892821
5	0.882979	0.866549	0.849174	0.855267
6	0.913121	0.904378	0.872848	0.885822
7	0.907801	0.885798	0.877338	0.880422
8	0.886525	0.856016	0.859351	0.857378
9	0.897163	0.887837	0.860811	0.865995
Rata-rata	0.900015	0.877913	0.872012	0.872863

Gambar 4. 33. Hasil Evaluasi *Cross Validation K-10 Decision Tree*



Gambar 4. 34. Hasil Evaluasi *Confusion Matrix Decision Tree*

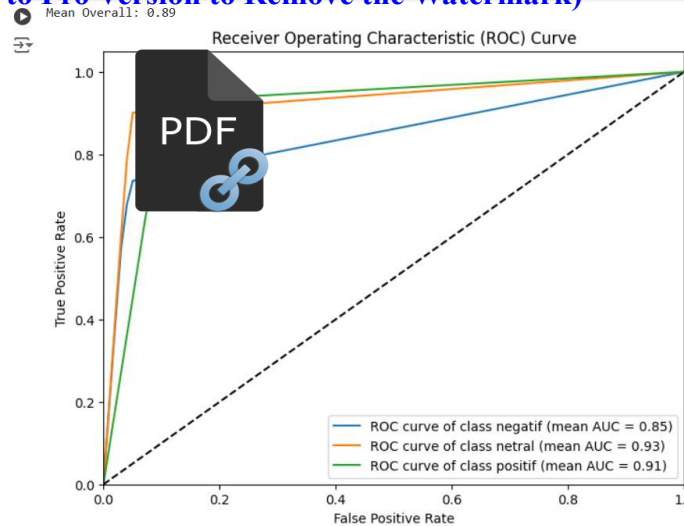
Overall Classification Report:

	precision	recall	f1-score	support
negatif	0.84	0.75	0.79	1041
netral	0.86	0.93	0.89	1230
positif	0.93	0.94	0.93	3370
accuracy			0.90	5641
macro avg	0.88	0.87	0.87	5641
weighted avg	0.90	0.90	0.90	5641

Gambar 4. 35. Hasil *Classification Report Decision Tree*

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



Gambar 4. 36. Grafik ROC AUC *Decision Tree*

4.3 Pembahasan

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap pilkada Jakarta 2024 pada media sosial twitter (x) menggunakan algoritma *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree*. Keenam algoritma tersebut akan dibandingkan performa mana yang terbaik dalam klasifikasi.

Penelitian ini dimulai dengan pengumpulan data (crawling) tweet menggunakan tweet-harvest dengan kata kunci “pilkada jakarta 2024” dimulai dari januari 2024-januari 2025 sebanyak 5.641 yang telah melalui tahapan preprocessing. Adapun hasil data yang sudah melewati tahapan *preprocessing* dapat dilihat sebagai berikut:

Tabel 4. 1. Data Setelah *Preprocessing*

Tweet	Hasil <i>Preprocessing</i>
Pilkada Jakarta 2024	5.641

Selanjutnya merupakan tahapan labeling menggunakan library *VaderSentiment*. Adapun hasil komposisi dari labeling dapat dilihat pada tabel berikut:

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Tabel 4. 2. Label Sentimen Vader

Tweet	Positif	Netral	Negatif
Pilkada Jakarta 2024	336	1230	1044

Tahap akhir yaitu penerapan klasifikasi menggunakan algoritma *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* dengan mengimplementasikan teknik *cross validation k-10* yang menghasilkan nilai rata-rata, evaluasi *confusion matrix* dan evaluasi ROC-AUC dapat dilihat sebagai berikut:

Tabel 4. 3. Hasil Klasifikasi SVM

Tweet	Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>Accuracy</i>	ROC AUC	CV K-10
Pilkada	Negatif	85%	80%	82%			
Jakarta	Netral	83%	91%	87%	90%	97%	90%
2024	Positif	95%	93%	94%			

Tabel 4. 4. Hasil Klasifikasi *Naïve Bayes*

Tweet	Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>Accuracy</i>	ROC AUC	CV K-10
Pilkada	Negatif	72%	70%	71%			
Jakarta	Netral	86%	48%	62%	79%	92%	79%
2024	Positif	79%	93%	86%			

Tabel 4. 5. Hasil Klasifikasi KNN

Tweet	Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>Accuracy</i>	ROC AUC	CV K-10
Pilkada	Negatif	79%	45%	57%			
Jakarta	Netral	46%	83%	59%	70%	83%	70%
2024	Positif	87%	73%	79%			

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Tabel 4. 6. Hasil Klasifikasi ANN

Tweet	Sentimen	Precision	Recall	F1 Score	Accuracy	ROC AUC	CV K-10
Pilkada	Negatif	74%	74%	79%			
Jakarta	Netral	82%	84%	83%	88%	96%	88%
2024	Positif	91%	94%	93%			

Tabel 4. 7. Hasil Klasifikasi *Random Forest*

Tweet	Sentimen	Precision	Recall	F1 Score	Accuracy	ROC AUC	CV K-10
Pilkada	Negatif	92%	61%	74%			
Jakarta	Netral	78%	79%	79%	85%	96%	85%
2024	Positif	86%	94%	90%			

Tabel 4. 8. Hasil Klasifikasi *Decision Tree*

Tweet	Sentimen	Precision	Recall	F1 Score	Accuracy	ROC AUC	CV K-10
Pilkada	Negatif	84%	75%	79%			
Jakarta	Netral	86%	93%	89%	90%	89%	90%
2024	Positif	93%	94%	93%			

Berdasarkan perbandingan antara keenam algoritma di atas, *Support Vector Machine* (SVM) muncul sebagai algoritma yang lebih baik dengan kinerja yang lebih tinggi. Dalam konteks ini, tweet pilkada Jakarta 2024 mencapai nilai precision untuk sentiment negatif sebesar 85% Sentiment netral sebesar 83% dan sentiment positif sebesar 95%. Pada nilai recall untuk sentiment negatif adalah 80%, sentiment netral adalah 91%, sentiment positif adalah 93%. Untuk nilai accuracy mencapai 90%, dan nilai ROC AUC sebesar 97% serta nilai rata-rata cross validation k-10 sebesar 90%.



5.1 Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa sentimen analisis terhadap pilkada Jakarta 2024 berdasarkan platform media sosial twitter (X) menggunakan Machine Learning, khususnya menggunakan algoritma *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest* dan *Decision Tree* yang menggunakan Data tweet dengan rentang waktu Januari 2024-Januari 2025 dengan kata kunci “pilkada jakarta 2024” menghasilkan 26.354 tweet. Setelah melewati tahapan preprocessing, menghasilkan data tweet sejumlah 5.641 tweet. Selanjutnya dilakukan labeling menggunakan VaderSentiment mendapatkan sentiment negatif sebanyak 1.044 data, sentimen netral 1.230 data dan sentimen positif sebanyak 3.367 data. Dari perbandingan Klasifikasi menggunakan enam algoritma, *Support Vector Machine* (SVM) menjadi algoritma yang lebih baik dengan kinerja yang lebih tinggi. Nilai precision untuk sentiment negatif sebesar 85% Sentiment netral sebesar 83% dan sentiment positif sebesar 95%. Pada nilai recall untuk sentiment negatif adalah 80%, sentiment netral adalah 91%, sentiment positif adalah 93%. Untuk nilai accuracy mencapai 90%, dan nilai ROC AUC sebesar 97% serta nilai rata-rata cross validation k-10 sebesar 90%. Dengan demikian, hasil dari penelitian ini memberikan pemahaman yang lebih baik tentang sentimen yang muncul di media sosial Twitter mengenai Pilkada Jakarta 2024 dan menunjukkan bahwa *Support Vector Machine* memiliki kinerja yang lebih baik dalam pengklasifikasian sentimen.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

5.2 Saran

Dari penelitian ini,  saran-saran untuk membantu peneliti selanjutnya, yaitu:

- a. Lakukan pembersihan data dengan baik agar data berkualitas, menghindari *noise* dan mendapatkan hasil yang optimal.
- b. Lakukan pengujian model terhadap isu-isu politik lain di Indonesia, misalnya Pilkada di provinsi lain, untuk mengukur generalisasi metode.
- c. Gunakan algoritma yang lebih cocok untuk dataset dalam jumlah besar, agar mendapatkan hasil yang lebih optimal, seperti algoritma BERT (Bidirectional Encoder Representations from Transformers).
- d. Untuk mengatasi data imbalance, bisa menggunakan SMOTE (Synthetic Minority Over-sampling Technique).

Demikian saran yang diharapkan menjadi dasar untuk penelitian selanjutnya yang lebih mendalam agar temuan penelitian ini bisa terus ditingkatkan untuk memberikan kontribusi yang lebih signifikan, baik di bidang akademik maupun praktis.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

DAFTAR PUSTAKA



- [1] U. Brawijaya, E. B. Sa... S. Hadi Wijoyo, And D. E. Ratnawati, “Fakultas Ilmu Komputer Analisis Sentimen Terhadap Pendapat Masyarakat Mengenai Pilkada 2024 Menggunakan Metode Support Vector Machine (Svm),” 2024. [Online]. Available: [Http://J-Ptiik.Ub.Ac.Id](http://j-ptiik.ub.ac.id)
- [2] R. Harun, R. Ishak, And S. S. Panna, “Analisis Sentimen Opini Publik Pengguna Twitter Terhadap Kenaikan Harga Bbm Menggunakan Algoritma Naïve Bayes,” *Copyright @Balok*, Vol. 2, No. 1, 2023.
- [3] A. Maulana, Inayah Khasnaputri Afifah, Asghafi Mubarrak, Kiagus Rachmat Fauzan, Ardhan Dwintara, And B. P. Zen, “Comparison Of Logistic Regression, Multinomialnb, Svm, And K-Nn Methods On Sentiment Analysis Of Gojek App Reviews On The Google Play Store,” *Jurnal Teknik Informatika (Jutif)*, Vol. 4, No. 6, Pp. 1487–1494, Dec. 2023, Doi: 10.52436/1.Jutif.2023.4.6.863.
- [4] M. N. Muttaqin And I. Kharisudin, “Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine Dan K Nearest Neighbor,” *Unnes Journal Of Mathematics*, Vol. 10, No. 2, Pp. 22–27, 2021, [Online]. Available: [Http://Journal.Unnes.Ac.Id/Sju/Index.Php/Ujm](http://journal.unnes.ac.id/sju/index.php/ujm)
- [5] D. Galih Pradana, M. L. Alghifari, M. Farhan Juna, And S. Dwisiwi Palaguna, “Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network,” *Indonesian Journal Of Data And Science (Ijodas)*, Vol. 3, No. 2, Pp. 55–60, 2022.
- [6] Yoga Religia, Agung Nugroho, And Wahyu Hadikristanto, “Klasifikasi Analisis Perbandingan Algoritma Optimasi Pada Random Forest Untuk Klasifikasi Data Bank Marketing,” *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, Vol. 5, No. 1, Pp. 187–192, Feb. 2021, Doi: 10.29207/Resti.V5i1.2813.
- [7] D. P. Sukma, S. Defit, And G. W. Nurcahyo, “Identifikasi Tingkat Kerusakan Peralatan Labor Teknik Komputer Jaringan Menggunakan Metode Decision Tree,” *Jurnal Sistim Informasi Dan Teknologi*, Pp. 275–280, Sep. 2021, Doi: 10.37034/Jsisfotek.V3i4.78.
- [8] A. Dwiki, A. Putra, And S. Juanita, “Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma Knn,” Vol. 8, No. 2, 2021, [Online]. Available: [Http://Jurnal.Mdp.Ac.Id](http://jurnal.mdp.ac.id)
- [9] J. Homepage *Et Al.*, “Malcom: Indonesian Journal Of Machine Learning And Computer Science Implementation Of Decision Tree Algorithm And Support

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- Vector Machine For Lung Cancer Classification Implementasi Algoritma Decision Tree Dan Support Vector Machine Untuk Klasifikasi Penyakit Kanker Paru,” Vol. 3, Pp. 2023.
- [10] E. Mardiani *Et Al.*, “Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regresi Terhadap Analisis Performa Pelajar Sma,” *Innovative: Journal Of Social Science Research*, Vol. 3, No. 2, Pp. 13880–13892, 2023, Accessed: Dec. 21, 2024. [Online]. Available: [Https://J-Innovative.Org/Index.Php/Innovative](https://J-Innovative.Org/Index.Php/Innovative)
- [11] A. Nata And S. Royal, “Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar,” 2022. [Online]. Available: [Http://Jurnal.Goretanpena.Com/Index.Php/Jssr](http://Jurnal.Goretanpena.Com/Index.Php/Jssr)
- [12] A. Mustofa And R. Novita, “Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter,” *Building Of Informatics, Technology And Science (Bits)*, Vol. 4, No. 1, Jun. 2022, Doi: 10.47065/Bits.V4i1.1628.
- [13] A. Triono, A. Setia Budi, And R. Abdillah, “Implementasi Peretasan Sandi Vigenere Chipher Menggunakan Bahasa Pemrograman Python,” 2023.
- [14] W. N. Mufidati Nur Edma, E. N. Andini, And I. Widodo, “Analisis Sentimen Pada Pengguna Tiktok Menggunakan Metode Random Forest (Studi Kasus: Jessica-Mirna),” *Journal Of Social Science Research*, Vol. 4, No. 3, Pp. 14477–14489, Jun. 2024, Doi: [Https://Doi.Org/10.31004/Innovative.V4i3.12240](https://doi.org/10.31004/Innovative.V4i3.12240).
- [15] F. Panjaitan, “Perbandingan Penggunaan Tfidfvectorizer, Countvectorizer, Dan Hashingvectorizer Dengan Optimalisasi Parameter Pada Machine Learning Untuk Analisis Sentimen Pemilu 2024,” *Jati (Jurnal Mahasiswa Teknik Informatika)*, Vol. 8, No. 4, Pp. 7413–7419, Aug. 2024.
- [16] A. Averina, H. Hadi, And J. Siswanto, “Analisis Sentimen Multi-Kelas Untuk Film Berbasis Teks Ulasan Menggunakan Model Regresi Logistik,” *Teknika*, Vol. 11, No. 2, Pp. 123–128, Jun. 2022, Doi: 10.34148/Teknika.V11i2.461.
- [17] D. Cahyanti, A. Rahmayani, And S. Ainy Husniar, “Indonesian Journal Of Data And Science Analisis Performa Metode Knn Pada Dataset Pasien Pengidap Kanker Payudara,” Vol. 1, No. 2, Pp. 39–43, 2020.
- [18] B. Zuraeni, J. Gajah Mada, K. Vitka City, And T. Ayu -Sekupang, “Ekosphere: Jurnal Ekonomi Pembangunan Dan Manajemen Volume Analisis Ramalan Cuaca Di Sekupang, Kota Batam Menggunakan Algoritma Decision Tree Dan Confusion Matrix,” *Ekosphere: Jurnal Ekonomi*

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- Pembangunan Dan Manajemen*, Vol. 1, No. 4, Pp. 15–26, Aug. 2024, [Online]. Available: <http://www.usinapublisher.org/index.php/ekosphere>
- [19] V. Fitriyana *Et Al.*, “Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine,” 2023.
- [20] H. Priatmojo, F. Saputra, M. H. Prasetyo, D. Puspitasari, And D. Nurlaela, “Perbandingan Klasifikasi Tingkat Penjualan Buah Di Supermarket Dengan Pendekatan Algoritma Decision Tree, Naive Bayes Dan K-Nearest Neighbor,” *Jurnal Insan (Journal Of Information Systems Management Innovation)*, Vol. 3, No. 1, 2023, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/jinsan>
- [21] R. Amanda *Et Al.*, “Analisis Sentimen Masyarakat Pada Teks Postingan Dan Komentar Di Platform X Dalam Pemilihan Presiden 2024 Di Indonesia Melalui Pemrograman Python Dan Metode Naive Bayes,” *Jurnal Membaca Bahasa & Sastra Indonesia*, 2024.
- [22] R. Wahyudi *Et Al.*, “Analisis Sentimen Pada Review Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine,” *Jurnal Informatika*, Vol. 8, No. 2, 2021, [Online]. Available: <http://ejournal.bsi.ac.id/ejournal/index.php/ji>
- [23] R. N. Mauliza And Y. R. Sipayung, “Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes,” *Technomedia Journal*, Vol. 9, No. 1, Pp. 1–16, Feb. 2024, Doi: 10.33050/Tmj.V9i1.2212.
- [24] A. Tangkelayuk And E. Mailoa, “Klasifikasi Kualitas Air Menggunakan Metode Knn, Naive Bayes Dan Decision Tree,” Vol. 9, No. 2, Pp. 1109–1119, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [25] D. D. Dewi, N. Qisthi, S. Sarah, S. Lestari, Z. Hidayah, And S. Putri, “Perbandingan Metode Neural Network Dan Support Vector Machine Dalam Klasifikasi Diagnosa Penyakit Diabetes,” *Jurnal Ilmiah Indonesia*, Vol. 3, No. 9, Pp. 828–839, 2023, Doi: 10.36418/Cerdika.Xxx.
- [26] K. Berfikir *Et Al.*, “Tarbiyah: Jurnal Ilmu Pendidikan Dan Pengajaran,” 2023. [Online]. Available: <https://jurnal.diklinko.id/index.php/tarbiyah/>
- [27] M. W. Magister, A. Pendidikan, U. Kristen, And S. Wacana, “Pendekatan Penelitian Pendidikan: Metode Penelitian Kualitatif, Metode Penelitian Kuantitatif Dan Metode Penelitian Kombinasi (Mixed Method),” Vol. 7, No. 1, Pp. 2896–2910, 2023.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- [28] F. Tangguh Admojo, "Indonesian Journal Of Data And Science Klasifikasi Aroma Alkohol Menggunakan Metode Knn," Vol. 1, No. 2, Pp. 34–38, 2020.
- [29] N. Puspitasari *Et Al.*, "Menerapkan Metode K-Nearest Neighbor Dan Fitur Warna Untuk Klasifikasi Daun Sirih Berdasarkan Citra Digital," Vol. 10, No. 2, 2023.
- [30] A. Syahril *Et Al.*, "Perbandingan Metode Decision Tree Dan K-Nearest Neighbor Terhadap Ulasan Pengguna Aplikasi Mypertamina Menggunakan Confusion Matrix," *Journal Of Information System Research (Josh)*, Vol. 5, No. 4, Pp. 1085–1094, 2024, Doi: 10.47065/Josh.V5i4.5639.
- [31] R. Indah Juwita Harahap, S. Khairani, And Corespondent Author, "Implementasi Metode K-Nearest Neighbor Untuk Klasifikasi Penyakit Tanaman Mentimun Pada Citra Daun Implementation Of The K-Nearest Neighbor Method For Classification Of Cucumber Diseases On Leaf Images," 2024. [Online]. Available: <https://jurnal.unity-academy.sch.id/index.php/jirsi/index>
- [32] F. Fatmawati And N. Narti, "Perbandingan Algoritma C4.5 Dan Naive Bayes Dalam Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Daring," *Jtim : Jurnal Teknologi Informasi Dan Multimedia*, Vol. 4, No. 1, Pp. 1–12, May 2022, Doi: 10.35746/Jtim.V4i1.196.
- [33] K. Kristiawan And A. Widjaja, "Perbandingan Algoritma Machine Learning Dalam Menilai Sebuah Lokasi Toko Ritel," *Jurnal Teknik Informatika Dan Sistem Informasi*, Vol. 7, No. 1, Apr. 2021, Doi: 10.28932/Jutisi.V7i1.3182.
- [34] A. Gaizka, A. R. Dzikrillah, And E. Sinduningrum, "Klik: Kajian Ilmiah Informatika Dan Komputer Analisis Sentimen Masyarakat Sebelum Dan Sesudah Terpilihnya Gibran Sebagai Cawapres Prabowo Menggunakan Naïve Bayes," *Media Online*, Vol. 4, No. 6, 2024, Doi: 10.30865/Klik.V4i6.1876.
- [35] I. Amelia *Et Al.*, "Analisis Sentimen Opini Publik Terhadap Pengambil Alihan Tmii Oleh Pemerintah Dengan Algoritma Naïve Bayes," Jul. 2023. [Online]. Available: <https://journals.upi-yai.ac.id/index.php/ikraith-informatika/issue/archive>
- [36] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, And W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *Jurnal Teknoinfo*, Vol. 14, No. 2, P. 115, Jul. 2020, Doi: 10.33365/Jti.V14i2.679.

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)




LAMPIRAN

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Lampiran 1. Form Pengajuan Judul



UNIVERSITAS BINA INSAN
Jalan Jenderal Besar H. H. S. R. Lubuklinggau Kec. Lubuklinggau Selatan I Kota Lubuklinggau Provinsi Sumatera Selatan

YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG

UNIVERSITAS BINA INSAN

Formulir Pengajuan Judul Skripsi
Program Studi Sistem Informasi


Nama : Biankha Ariesty
NIM : 2102030039
Alamat : Jl. Air Temam. Perumahan Griya Air Temam Blok C.11
No.Hp : 083801164741

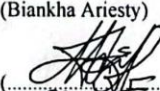
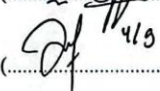
Rumusan Masalah 1 : Bagaimana Melakukan Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Opini X Menggunakan Machine Learning?
Judul 1 ✓ : Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Opini X Menggunakan Machine Learning

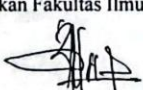
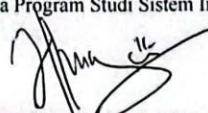
Rumusan Masalah 2 : Bagaimana Penerapan Analisis Sentimen Pada Ulasan Aplikasi My Icon Plus di Google Play Store Menggunakan Machine Learning?
Judul 2 : Penerapan Analisis Sentimen Pada Ulasan Aplikasi My Icon Plus di Google Play Store Menggunakan Machine Learning

Rumusan Masalah 3 : Bagaimana Menerapkan Metode Naive Bayes Dan SVM Untuk Analisis Sentimen Terhadap Aplikasi Youtube Kids Berdasarkan Ulasan Pengguna di Google Play Store?
Judul 3 : Analisis Sentimen Terhadap Aplikasi Youtube Kids Berdasarkan Ulasan Pengguna di Google Play Store Menggunakan Metode Naive Bayes Dan SVM.

Diusulkan Judul Nomor 1(satu) 2(Dua) 3(Tiga)*

Lubuklinggau,.....2024
Mahasiswa yang mengusulkan,

(Biankha Ariesty)

Menyetujui Dosen Pembimbing,
Pembimbing 1 (Muhammad Nur Alamsyah, M.Kom) ( 5/9 2024)
Pembimbing 2 (Nelly Khairani Daulay, M.Kom) ( 4/9 2024)

Mengesahkan
Dekan Fakultas Ilmu Teknik Mengetahui
Ketua Program Studi Sistem Informasi
 
(Dr. Rudi Kurniawan, S.T., M.Kom) (Harma Oktafia Lingga Wijaya, M.Kom)

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Lampiran 2. SK Pembimbing

No	NIM	Nama Mahasiswa	Pembimbing 1	Pembimbing 2
31	2102030038	Regina Rizki Putri	Budi Santoso, M. Kom	Davit Irawan, M. Kom
32	2102030039	Biankha Ariesty	Muhammad Nur Alamsyah, M. Kom	Nelly Khairani Daulay, M. Kom
33	2102030041	Antika Yupi Yolanda	Bunga Intan, M. Kom	Ahmad Sobri, M. Kom
34	2102030042	Anggi Mika	Lukman Hakim, M. Kom	Satrianansyah, M. Kom
35	2102030044	Cindy Eryan Lorenza	Dr. M. Agus Syamsul Arifin, S. St., M. Kom	Cindi Wulandari, M. Kom
36	2102030045	Rahma Sri Utami	Dr. Susanto, M. Kom	Lukman Hakim, M. Kom
37	2102030048	Fysca Arlisty	Dr. Rudi Kurniawan, S.T., M. Kom	Andri Anto Tri Susilo, M. Kom
38	2202030041p	Nanda Cornelis Rasyid	Joni Karman, M. Kom	Asep Toyib Hidayat, M. Kom
39	2302030057P	Henny Yuniarti Rabeka	Asep Toyib Hidayat, M. Kom	Muhammad Nur Alamsyah, M. Kom

Ditetapkan di : Lubuklinggau
 Pada tanggal : 29 Agustus 2024
 Dekan Fakultas Ilmu Teknik,



Dr. Rudi Kurniawan, S.T., M. Kom

Tembusan Yth:
 1. Ketua Yayasan Pendidikan Dwi Tunggal Palembang (sebagai laporan)
 2. Rektor Universitas Bina Insan (sebagai laporan)
 3. Arsip

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Lampiran 3. SK Penguji



UNIVERSITAS BINA INSAN
FAKULTAS ILMU TEKNIK
Lubuklinggau, Kalimantan Selatan

KEPUTUSAN
FAKULTAS ILMU TEKNIK
Nomor : 1025 / UNIV.BI/ FIT.I/SK/2024

TENTANG
PENGANGKATAN DOSEN PENGUJI PROPOSAL SKRIPSI MAHASISWA
PROGRAM STUDI SISTEM INFORMASI FAKULTAS ILMU TEKNIK
UNIVERSITAS BINA INSAN LUBUKLINGGAU

DENGAN RAHMAT TUHAN YANG MAHA ESA, UNIVERSITAS BINA INSAN LUBUKLINGGAU

Memperhatikan : Bahwa dengan selesainya mahasiswa menyusun Proposal Skripsi Program Studi Sistem Informasi Fakultas Ilmu Teknik Tahun Akademik 2024/2025, maka perlu menunjuk dan mengangkat Dosen Penguji Proposal Skripsi untuk menguji mahasiswa tersebut dalam menyelesaikan kuliahnya di lingkungan Universitas Bina Insan Lubuklinggau;

Menimbang : 1. Bahwa dalam upaya menyelenggarakan pendidikan tinggi yang berkualitas dipandang perlu mengangkat Dosen Penguji Proposal Skripsi di lingkungan Universitas Bina Insan Lubuklinggau;
2. Sehubungan dengan butir 1 (satu) tersebut di atas, maka dipandang perlu mengeluarkan Surat Keputusan sebagai landasan hukumnya;

Mengingat : 1. Undang-Undang No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional;
2. Peraturan Pemerintah Republik Indonesia No. 60 Tahun 1999 tentang Pendidikan Tinggi;
3. Keputusan Menteri Pendidikan Nasional Republik Indonesia No. 232/U/2000 tentang Pedoman Penyusunan Kurikulum Pendidikan Tinggi dan Penilaian Hasil Belajar Mahasiswa;
4. Keputusan Menteri Pendidikan Nasional Republik Indonesia No. 184/U/2001 tentang Pedoman Pengawasan-pengendalian dan Pembinaan Program Diploma, Sarjana dan Pascasarjana di Perguruan Tinggi;
5. SK Menteri Riset, Teknologi Dan Pendidikan Tinggi Republik Indonesia Nomor 223/KPT/A/2019 Tentang Izin Penggabungan Sekolah Tinggi Ilmu Ekonomi Musi Rawas dan Sekolah Tinggi Manajemen dan Ilmu Komputer Musi Rawas Menjadi Universitas Bina Insan;
6. SK Ketua Yayasan Nomor 01.113/YPDT-Pjg/KP/SK/IV/2019 Tentang Pengangkatan Rektor Universitas Bina Insan Lubuklinggau;
7. SK Rektor Universitas Bina Insan Nomor 1235/UNIV.BI/R/KP/SK/2020 Tentang Pengangkatan Pejabat Pada Universitas Bina Insan Lubuklinggau;
8. Statuta Universitas Bina Insan Lubuklinggau;

MEMUTUSKAN

Mentapkan Pertama : Mengangkat Saudara yang namanya tercantum pada lampiran ini, sebagai penguji Proposal Skripsi Program Studi Sistem Informasi Fakultas Ilmu Teknik Tahun Akademik 2024/2025 di Universitas Bina Insan Lubuklinggau;


Kedua : Semua biaya yang timbul akibat dikeluarkannya Surat Keputusan ini dibebankan kepada anggaran Universitas Bina Insan Lubuklinggau atau dana khusus yang disediakan untuk itu;

Ketiga : Kepada yang bersangkutan diberikan honorarium yang besarnya sesuai dengan peraturan Universitas Bina Insan Lubuklinggau;

Keempat : Surat Keputusan ini berlaku sejak tanggal ditetapkan, dengan ketentuan apabila ternyata dikemudian hari terdapat kekeliruan dalam penetapan surat keputusan ini, akan diperbaiki sebagaimana mestinya.

Demikian Surat Keputusan ini ditetapkan untuk dilaksanakan sebagaimana mestinya.

Ditetapkan di : Lubuklinggau
Pada tanggal : 11 Desember 2024
Dekan Fakultas Ilmu Teknik




Dr. Rudi Kurniawan, S.T., M.Kom

Tembusan Yb:

1. Ketua Yayasan Pendidikan Dwi Tunggal Palembang (sebagai laporan)
2. Rektor Universitas Bina Insan (sebagai laporan)
3. Arsip

☎ 0733-4553932 (Rektorat Universitas) 📠 0812-3826-6228 (Marketing UNIVBI)
 ☎ 0733-3280300 (Bina Insan) 📠 0652-3151-5800 (Admin UNIVBI)
 ☎ 0733-3280200 (Pascasarjana) 📠 Admin@univbinainstan.ac.id 🌐 univbinainstan.ac.id / pasca.univbinainstan.ac.id

Dipindai dengan


Lampiran Surat Keputusan Dekan Fakultas Ilmu Teknik Universitas Bina Insan Lubuklinggau
Nomor : 1025 / UNIV.BI/ FIT.I/SK/2024
Tanggal : 11 Desember 2024
Tentang : Susunan Pengangkatan Dosen Penguji Sidang Proposal Skripsi Program Studi Sistem Informasi IA. 2024/2025

No	Nama Mahasiswa	NIM	Ketua	Sekretaris	Anggota	Hari	Tanggal	Jam	Ruang
1	Nabila	210203008	Dr. Rudi Kurniawan, ST., M.Kom	Joni Karman, M.Kom	Muhammad Nur Alamsyah, M.Kom	Jun'at	13/12/2024	08.00-09.00	Ruang Sidang 1
2	Detha Ayu Arini	210203013	Bunga Intan, M.Kom	Muhammad Irvai, M.Kom	Dr. Rudi Kurniawan, ST., M.Kom	Jun'at	13/12/2024	09.00-10.00	Ruang Sidang 1
3	Fysca Artisty	210203048	Dr. Rudi Kurniawan, ST., M.Kom	Andri Anto Tri Susilo, M.Kom	Bunga Intan, M.Kom	Jun'at	13/12/2024	10.00-11.00	Ruang Sidang 1
4	Biankha Ariesty	210203039	Muhammad Nur Alamsyah, M.Kom	Nelly Khairani Dauloy, M.Kom	Dr. Rudi Kurniawan, ST., M.Kom	Jun'at	13/12/2024	11.00-12.00	Ruang Sidang 1

Lubuklinggau, 11 Desember 2024
Dekan Fakultas Ilmu Teknik



Dr. Rudi Kurniawan, S.T., M.Kom
FAKULTAS ILMU TEKNIK

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Lampiran 4. Dataset Awal

	Source Name	id_str	image_url_in_reply_tiang	location	quote
1	Source:NaI conversati created_ a favorite_c full_text				
2	1.87E+18 Thu Dec 1	0 Ridwan Kamil Gagal Total di Pilkada Jakarta Jhon Sitorus Ungkap Kegagalan Tim Suksesnya https://t.1.87E+18	in	Bengkulu	
3	1.87E+18 Thu Dec 1	47 Kegagalan kubu sebelah dalam mengajukan gugatan ke MK menunjukkan bahwa kemenangan Pramn https://pbs.twimg.com/1.87E+18	in	Indonesia	
4	1.87E+18 Sun Dec 0	7 @aewin86 Bujer-bujer Hasto sekarang pada orgasme euforia Pramono menang di pilkada Jakarta. Et https://pbs.twimg.com/1.87E+18	in		
5	1.86E+18 Tue Dec 0	20 Tim Pemenangan Pasangan Calon Ridwan Kamil-Suswono (RIDO) mengatakan pelaksanaan Pilkada J https://pbs.twimg.com/1.86E+18	in	Pulo Gadu	
6	1.86E+18 Thu Nov 2	10 Potret nyata kegagalan parpol dalam mengemban aspirasi masyarakat sehingga memaksa masyarakat https://pbs.twimg.com/1.86E+18	in		
7	1.85E+18 Wed Oct 3	97 Dari ramainya seteru pilkada Jateng & Jakarta. Saya jadi paham ada irisan pendukung. Mana pe https://pbs.twimg.com/1.85E+18	in	Indonesia	
	1.84E+18 Fri Oct 25	101 Kata kader PKS Suswono - Singapura Ketar-Ketir jika Ridwan Kamil Menang Pilkada. Padahal warga Jah https://pbs.twimg.com/1.84E+18	in		
	1.84E+18 Tue Oct 0	0 Kekanak-kanakan Pihak yang mengkorifikasi kegagalan Anies karena kekuatan politik tertentu sesun https://pbs.twimg.com/1.84E+18	in		
	1.84E+18 Sun Oct 0	3 Bukti Kegagalan RK Dalam Rancangan Halte Di Bandung Akan Memriku Kegagalannya Di Pilkada Jakart https://pbs.twimg.com/1.84E+18	in	Burni Man	
	1.84E+18 Thu Sep 2	2 @FelixSGL1 Yap. Tapi yang lumayan mengemukakan adalah suara Syaikh yang cuma mentok dibawah https://pbs.twimg.com/1.84E+18	in		
	1.84E+18 Wed Sep 2	0 Jika di Pilkada Jakarta angka golput meningkat / coblos semua paslon gub meningkat itu artinya sebu https://pbs.twimg.com/1.84E+18	in	Indonesia	
12	1.83E+18 Tue Sep 0	3 Kegagalan Anies maju pilkada Jakarta 2024 murni di tangan ketua umum parpol Pope Francis Anya Sa https://pbs.twimg.com/1.83E+18	in	Bandung	
13	1.83E+18 Tue Sep 0	1 Respon Partai Gerindra Enggan Mencampuri Kegagalan Anies Terkait Pilkada Jakarta https://t.co/OFH.1.83E+18	in		
14	1.83E+18 Sun Sep 0	0 Jawaban Anies Dianya Dukung Ridwan Kamil atau Pramono Anung di Pilkada Jakarta 2024 https://t.1.83E+18	in		
15	1.83E+18 Sat Aug 31	1 HASHTAG Tagar populer di publik adalah #AniesBaswedan #Pilkada dan #Pilpres menunjukkan fokus https://pbs.twimg.com/1.83E+18	in	Jakarta Se	
16	1.83E+18 Sat Aug 31	1 https://t.co/c4lDQKtYpj Pendiri Rumah Demokrasi Ramdaniyah mengungkapkan bahwa kegagalan / https://pbs.twimg.com/1.83E+18	in	Jalan Cier	
17	1.83E+18 Fri Aug 30	0 @tfssembiring PKS akan ditnggal sama pendukungnya dengan kegagalan Anis maju pilkada Jakarta https://pbs.twimg.com/1.83E+18	in		
18	1.83E+18 Fri Aug 30	1 @cakiminNOW Karena coatail effect Anies Baswedan PGB bertambah kursi diparlemen berbeda deng https://pbs.twimg.com/1.82E+18	in	Indonesia	
19	1.82E+18 Mon Aug :	1 @CinereUnun41935 @nanath27 @Kimberley_P508 Sudah diblaingin fokus saja bantu pilkada Jakarta https://pbs.twimg.com/1.82E+18	in	West Borr	
20	1.82E+18 Tue Jul 30				
21	1.82E+18 Wed Jul 2				

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)


Lampiran 5. Dataset Setelah Preprocessing

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V				
1	full text	cleansing	case	foldl	tokenize	stopword	normalisa	english	l	stemmed	tweet												
2	Ridwan KaRidwan	KaRidwan	ka	['ridwan',	['ridwan',	['ridwan',	['ridwan',	['ridwan',	['kamil',	['tail',	['total',	['pilkada',	['jakarta',	['jhon',	['storu',	['failur',	['team',	['success',]					
3	Kegagalan	Kegagalan	kegagalan	['kegagala	['kegagala	['kegagala	['failure',	['failur',	['camp',	['next',	['subomi',	['lawsuit',	['mk',	['victori',	['pramono',	['anung',	['trano',	['karno',	['pilkada',	['jakarta',			
4	@aewin8t	Bujerbujie	bujerbujie	['bujerbuj	['bujerbuj	['bujerbuj	['bujerbuj',	['bujerbuj',	['hasto',	['orgasm',	['euphoria',	['pramono',	['win',	['pilkada',	['jakarta',	['euphoria',	['cover',	['failur',	['jateng',	['bar			
5	Tim Peme	Tim Peme	tim peme	['tim',	['per['tim',	['per['team',	['w['team',	['winner',	['pair',	['candid',	['ridwan',	['kamiluswono',	['rdo',	['implement',	['pilkada',	['jakarta',	['highlight',	['sharp',					
6	Potret nva	Potret nva	potret nva	['potret',	['['potret',	['['portrait',	['['portrait',	['real',	['failur',	['polit parti',	['carri out',	['aspir',	['societi',	['forc',	['societi',	['choos',	['figur',	['candid',	['leader',	['acc			
7	rana dari	rana dari	rana	['rani',	['dari',	['rai['ramainy',	['ramainy['crowd',	['['crowd',	['enemi',	['pilkada',	['jateng',	['amp',	['jakarta',	['understand',	['fisan',	['support',	['support',	['prabowo',	['projo',	['giti			
8	bnakk	bnakk	bnakk	['kekanak',	['kekanak',	['kekanak',	['childish',	['childish',	['glorifi',	['failur',	['ani',	['power',	['polit',	['reali',	['act',	['childish',	['failur',	['ani',	['pilkada',	['jakarta',	['comprehens		
9	kade	kade	kade	['kata',	['ka',	['kader',	['f['kader',	['p['cadr',	['pk',	['suswono',	['singapura',	['ketarketr',	['ridwan',	['kamil',	['win',	['pilkada',	['citizen',	['jakarta',	['yang',	['ketar',	['ke		
10	bnakk	bnakk	bnakk	['kekanak',	['kekanak',	['kekanak',	['childish',	['childish',	['glorifi',	['failur',	['ani',	['power',	['polit',	['reali',	['act',	['childish',	['failur',	['ani',	['pilkada',	['jakarta',	['comprehens		
11	@FelixSGI	Yap Tapi	yap tapi	['yap',	['tap',	['yap',	['lun['yap',	['lun['yap',	['noi',	['yap',	['notbad',	['surpris',	['sound',	['svaikhu',	['stuck',	['under',	['jabar',	['wrong',	['lumbang',	['sound',	['pk',	['thought',	['mo
12	Jika di	pilk	jika di	['pilk',	['di',	['pilkada',	['['pilkada',	['['pilkada',	['jakarta',	['number',	['golput',	['increas',	['vote',	['pasion',	['gub',	['inreas',	['resist',	['peopl',	['to',	['polit parti',	['wh		
13	Kegagalan	Kegagalan	kegagalan	['kegagala	['kegagala	['kegagala	['failure',	['failur',	['ani',	['forward',	['pilkada',	['jakarta',	['pure',	['hand',	['chairman',	['polit parti',	['pope',	['franci',	['arya',	['saloka',	['sp		
14	Respon	Respon	respon	['respon',	['respon',	['respon',	['response',	['respons',	['parti',	['gerindra',	['relect',	['interf',	['failur',	['ani',	['relat',	['pilkada',	['jakarta',	['new',]					
15	Jawaban	Jawaban	jawaban	['jawaban',	['awies',	['d',	['awies',	['s',	['ani',	['support',	['ridwan',	['kamil',	['pramono',	['anung',	['pilkada',	['jakarta',	['lesson',	['soul',	['preciou',	['to',	['judg',	['put dc	
16	HASHTAG	HASHTAG	hashtag	['hashtag',	['hashtag',	['hashtag',	['hashtag',	['home',	['democraci',	['ramdansyah',	['failur',	['ani',	['baswedan',	['presidenti elect',	['focu',	['nomin',	['ani',	['cathmulyor					
17	https://t.c	Pendiri	pendiri	['pendiri',	['pendiri',	['founder',	['founder',	['home',	['democraci',	['ramdansyah',	['failur',	['ani',	['baswedan',	['presidenti elect',	['focu',	['nomin',	['ani',	['cathmulyor					
18	https://t.c	Kegagala	kegagala	['kegagala',	['kegagala',	['kegagala',	['failure',	['failur',	['ani',	['vehicl',	['polit',	['forward',	['pilkada',	['jakarta',	['hope for',	['associ',	['palac',	['aniesbaswedan',	['pilkada',	['pre			
19	@titsembi	PKS akan	pk akan	['pks',	['akal',	['pks',	['diti',	['pks',	['diti',	['pk',	['left behind',	['support',	['failur',	['ani',	['forward',	['pilkada',	['jakarta',]						
20	@cakimin	Karena cc	karena cc	['karena',	['coattail',	['coattail',	['coattail',	['coattail',	['coattail',	['coattail',	['effect',	['ani',	['baswedan',	['pkb',	['increas',	['seat',	['in parliament',	['differ',	['ppp',	['consequ',	['wrong',	['support	
21	@CinereU	Sudah di	sudah di	['sudah',	['diblangiri',	['diblangiri',	['diblangiri',	['diblangiri',	['diblangiri',	['diblangiri',	['focu',	['help',	['pilkada',	['jakarta',	['failur',	['take care of',	['hous',	['slum',	['peopl',	['poor',	['leadership',	['si',	['tr'

Protected by PDF Anti-Copy Free

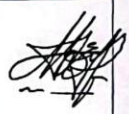

(Upgrade to Pro Version to Remove the Watermark)

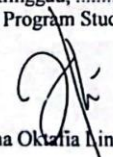
Lampiran 7. Lembar Bimbingan P1


PDF

LEMBAR BIMBINGAN SKRIPSI

Nama : Biankha Ariesty
 Nim : 2102030039
 Program Studi : Sistem Informasi
 Pembimbing 1 : Muhammad Nur Alamsyah, M.Kom
 Pembimbing 2 : Nelly Khairani Daulay, M.Kom
 Judul : Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
	20-12-2024		-Perbaiki hal 44-45 -jelaskan dari tiap proses gambar yang dilakukan.		
	4-1-2025		ACC ujian		

Lubuklinggau,2024
 Ketua Program Studi Sistem Informasi

 (Harma Oktaria Lingga Wijaya, M.Kom)

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Lampiran 8. Lembar Bimbingan P2

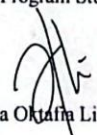


LEMBAR Bimbingan SKRIPSI

Nama : Biankha Ariesty
 Nim : 2102030039
 Program Studi : Sistem Informasi
 Pembimbing 1 : Muhammad Nur Alamsyah, M.Kom
 Pembimbing 2 : Nelly Khairani Daulay, M.Kom
 Judul : Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
1.	14-12-2024	Bab I	perbaiki urut kelaseay		<i>[Signature]</i>
2.	16-12-2024	Bab II	penelitian relikah delaskan suda permasalahan yg diteliti, metodanya serta hasilnye		<i>[Signature]</i>
3.	17-12-2024	Bab III	- Pada tabel dan gambar larisca keterangan dan kaitkan nomor nomornya. - Setiap penulisan rumus berikan nomor		<i>[Signature]</i>
4.	18-12-2024	Bab IV	Perhatikan lagi hasil Evaluasi		<i>[Signature]</i>
5.	19-12-2024	Bab V	kesimpulan dibuat paragraf suda. tidak point perpoint		<i>[Signature]</i>
6.			ACC Lanjut pembimbing # I		


Lubuklinggau,2024
Ketua Program Studi Sistem Informasi


 (Harma Oktifa Lingga Wijaya, M.Kom)

Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)







Lampiran 9. Lembar Perbaikan Ujian Skripsi



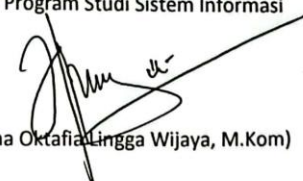
UNIVERSITAS BINA INSAN
Jalan Jendral Besar
Lubuk Kumpang Kec. Lubuklinggau Selatan I Kota Lubuklinggau Provinsi Sumatera Selatan

LEMBAR PERBAIKAN UJIAN SKRIPSI

Nama Mahasiswa : Biankha Ariesty
 NIM : 2102030039
 Jenjang Pendidikan : Strata 1 (S1)
 Fakultas : Ilmu Teknik
 Program Studi : Sistem Informasi
 Konsentrasi : Data Science
 Judul : Sentimen Analisis Terhadap Pilkada Jakarta 2024 Berdasarkan Platform Media Sosial Twitter (X) Menggunakan Machine Learning

No	Dosen Penguji	Komentar Perbaikan	Tanda Tangan Ujian	Tanda Tangan Revisi
1	Muhammad Nur Alamsyah, M.Kom	Perbaiki Tata tulis		
2	Nelly Khairani Daulay, M.Kom	Bagian kesimpulan buat dalam bentuk 1 Paragraf		
3	Dr.Rudi Kurniawan, S.T.,M.Kom	Bagian Saran ditambahkan Solusi Penanganan Data Imbalance		

Lubuklinggau,2025
 Ketua Program Studi Sistem Informasi


 (Harma Oktavia Lingga Wijaya, M.Kom)

0733-4553932 (Rektorat Universitas) 0812-1826-6228 (Marketing UNIVBI)
 0733-3280300 Bina Insan 0852-3151-5800 (Admin UNIVBI)
 0733-3280200 (Pascasarjana) Admin@univbinainsan.ac.id univbinainsan.ac.id - pasca.univbinainsan.ac.id