

TESIS
ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI
TWITTER DENGAN MACHINE LEARNING



disusun oleh

ARIF DWI SYAFUTRA

22.55.2298

Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

TESIS
ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI
TWITTER DENGAN MACHINE LEARNING

SENTIMENT ANALYSIS ABOUT OMNIBUS LAW ON
TWITTER WITH MACHINE LEARNING

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana
Program Studi S2 Informatika



disusun oleh

ARIF DWI SYAFUTRA

22.55.2298

Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2026

HALAMAN PERSETUJUAN

**ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI TWITTER
DENGAN MACHINE LEARNING**

**SENTIMENT ANALYSIS ABOUT OMNIBUS LAW ON TWITTER WITH
MACHINE LEARNING**

yang disusun dan diajukan oleh

Arif Dwi Syafutra

22.55.2298

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 01 Desember 2025

Dosen Pembimbing,



Prof. Dr. Kusrini, M.Kom

NIK. 190302106

HALAMAN PENGESAHAN

**ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI TWITTER
DENGAN MACHINE LEARNING**

**SENTIMENT ANALYSIS ABOUT OMNIBUS LAW ON TWITTER WITH
MACHINE LEARNING**

yang disusun dan diajukan oleh

Arif Dwi Syafutra

22.55.2298

Telah dipertahankan di depan Dewan Penguji
pada tanggal 01 Desember 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024



Emha Taufik Luthfi, S.T., M.Kom.
NIK. 190302125



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 01 Desember 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Arif Dwi Syafutra
NIM : 22.55.2298
Konsentrasi : Business Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
**ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI TWITTER
DENGAN MACHINE LEARNING**

Dosen Pembimbing Utama : Prof. Dr. Kusriani, M.Kom.
Dosen Pembimbing Pendamping : -

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 01 Desember 2025
Yang Menyatakan,



Arif Dwi Syafutra

HALAMAN PERSEMBAHAN

Bismillahirrahmanirrahim

Dengan rasa syukur yang tak terhingga, penelitian tesis ini saya persembahkan sebagai ungkapan terima kasih kepada Allah SWT, Yang Maha Pengasih dan Penyayang, atas segala rahmat, petunjuk, dan karunia-Nya yang melimpah dalam setiap langkah perjalanan hidup.

Kepada orang tua tercinta, yang telah memberikan cinta, doa, dan dukungan tak terhingga selama ini. Terima kasih atas pengorbanan, kasih sayang, dan dorongan yang memotivasi untuk terus berjuang dan berkarya. Kepada istri tercinta, yang selalu setia mendampingi, memberikan semangat, dan menjadi penyejuk hati di setiap suka dan duka. Terima kasih atas kesabaran, pengertian, dan doa yang tak pernah putus. Kepada dosen pembimbing dan penguji yang terhormat, atas bimbingan, ilmu, dan kesabaran dalam mengarahkan penelitian ini hingga selesai dengan baik. Kepada seluruh jajaran Dinas Komunikasi dan Informatika Kota Samarinda serta rekan-rekan kantor, atas dukungan, kerjasama, dan semangat yang telah diberikan.

Tidak lupa kepada Universitas AMIKOM Yogyakarta, atas fasilitas, ilmu pengetahuan, dan pengalaman yang telah diberikan selama proses pendidikan. Terima kasih kepada seluruh dosen, staf, dan rekan-rekan yang telah memberikan inspirasi dan bantuan. Semoga segala jerih payah, doa, dan keikhlasan yang telah diberikan menjadi amal jariyah yang bermanfaat bagi kita semua.

HALAMAN MOTTO

"Allah akan meninggikan orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu pengetahuan beberapa derajat. Dan Allah Maha Mengetahui apa yang kamu kerjakan, Ya Tuhanku, kasihanilah keduanya (ibu bapakku), sebagaimana mereka berdua telah mendidik aku pada waktu kecil."

QS. Al-Mujadalah: 11 & QS. Al-Isra: 24



KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis ini yang merupakan syarat dalam menyelesaikan jenjang Pendidikan S2 Pascasarjana Teknik Informatika Universitas AMIKOM Yogyakarta yang berjudul "ANALISIS SENTIMEN TENTANG OMNIBUS LAW DI TWITTER DENGAN MACHINE LEARNING" dengan baik.

Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Kedua orang tua tercinta, Bapak Suwarno dan Ibu Musiroh, atas doa, pengorbanan, dan dukungan yang tak terhingga selama ini.
2. Istri tercinta, Ayuni Sukarman Lokut, yang selalu setia mendampingi, memberikan semangat, dan doa dalam setiap langkah perjalanan pendidikan ini.
3. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
4. Ibu Prof. Dr. Kusriani, M.Kom., selaku dosen pembimbing utama yang telah banyak membantu memberikan ilmu, bimbingan, masukan, dan saran dalam penelitian ini.
5. Ibu Prof. Dr. Ema Utami, S.Si., M.Kom., selaku Direktur Pascasarjana Universitas AMIKOM Yogyakarta.
6. Dosen pembimbing pendamping yang telah mengarahkan penulis dan memberikan saran yang membangun dalam penelitian ini.

7. Segenap dosen penguji yang telah memberikan masukan dan arahan untuk perbaikan tesis ini.
8. Segenap dosen dan staf Magister Teknik Informatika Universitas AMIKOM Yogyakarta yang telah memberikan ilmu, wawasan, bantuan, dan pengalaman baru pada penulis selama perkuliahan.
9. Rekan-rekan kantor di Dinas Komunikasi dan Informatika Kota Samarinda: Alfi, Yasin, Eka Septy, Aulia, Putri, Naya, dan Gary, yang telah memberikan dukungan, kerjasama, dan semangat.
10. Rekan-rekan seperjuangan MTI Konsentrasi Business Intelligence yang memberikan motivasi, pengalaman baru, dan semangat yang luar biasa semasa perkuliahan dan penyelesaian tesis ini.

Penulis menyadari bahwa dalam penyusunan tesis ini masih terdapat kekurangan. Oleh karena itu, penulis dengan sangat senang hati menerima kritik dan saran yang membangun dari pembaca. Akhir kata, penulis berharap semoga tesis ini dapat memberikan manfaat bagi yang membacanya.

Yogyakarta, 02 Desember 2025

Penulis

DAFTAR ISI

HALAMAN PENGESAHAN	Error! Bookmark not defined.
HALAMAN PERSETUJUAN	v
HaLAMAN PERNYATAAN KEASLIAN TESIS	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
kaTA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvii
INTISARI	xviii
<i>ABSTRACT</i>	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 Tinjauan Pustaka	6
2.2 Landasan Teori	8
2.2.1 <i>Natural Language Processing (NLP)</i>	8
2.2.2 Analisis Sentimen	8
2.2.3 Twitter	10
2.2.4 Emosi	11

2.2.5	<i>Pre-Processing</i>	11
2.2.6	<i>Feature Extraction</i>	14
2.2.7	<i>Lexycon Based Sentiment Analysis</i>	15
2.2.8	<i>Support Vector Machine (SVM)</i>	17
2.2.9	<i>Random Forest</i>	18
2.2.10	<i>Multinomial Naïve Bayes</i>	19
2.2.11	<i>SMOTE (Synthetic Minority Over-sampling Technique)</i>	19
2.2.12	<i>Resampling dengan Penyeimbangan Ulang (Under-sampling dan Over-sampling)</i>	20
2.2.13	<i>Evaluasi Model</i>	22
2.3	Keaslian Penelitian.....	25
BAB III Metode Penelitian		36
3.1	Metode Pengumpulan Data.....	36
3.2	Metode Analisis Data.....	36
3.3	Jenis, Sifat dan Pendekatan Penelitian.....	38
3.4	38
BAB IV HASIL DAN PEMBAHASAN		45
4.1	Pengumpulan dan Karakteristik Data.....	45
4.1.1	Deskripsi data.....	45
4.1.2	Distribusi data per tahun & per bulan (volume tweet dan rasio positif-negatif).....	47
4.1.3	Deskriptif dataset + tabel distribusi label awal.....	50
4.2	Pra-Pemrosesan Data.....	51
4.2.1	<i>Case folding dan cleaning</i>	51

4.2.2	Normalisasi kata tidak baku → kata baku.....	54
4.2.3	Tokenisasi & stopwords removal (daftar kata yang dihapus/dipertahankan).....	57
4.2.4	Stemming menggunakan Sastrawi.....	59
4.2.5	Rekap tahapan preprocessing dan hasil akhir data bersih.....	62
4.2.6	Dampak preprocessing terhadap volume data dan kejelasan teks.....	64
4.3	Pelabelan Sentimen & Eksplorasi Data.....	67
4.3.1	Pelabelan sentimen dengan lexicon-based.....	67
4.3.2	Visualisasi distribusi sentimen positif-negatif.....	69
4.3.3	<i>Word cloud</i> & analisis kata dominan.....	71
4.3.4	Interpretasi hubungan kata dominan dengan konteks sosial (buruh, investasi, hukum).....	74
4.4	Ekstraksi dan Seleksi Fitur.....	77
4.4.1	Penerapan TF-IDF (unigram–bigram–trigram).....	77
4.4.2	Penambahan <i>smart features</i> (word_count, positive_count, dll).....	83
4.4.3	Seleksi fitur (Chi-square, korelasi, dan dampaknya).....	86
4.4.4	Tabel hasil reduksi fitur dan performa model awal.....	89
4.5	Pembentukan dan Evaluasi Model.....	92
4.5.1	Desain eksperimen dan pembagian data.....	92
4.5.2	Hasil model SVM Linear (tanpa resampling).....	98
4.5.3	Hasil model <i>Random Forest</i> (tanpa resampling).....	101
4.5.4	Hasil model Multinomial Naïve Bayes (tanpa resampling).....	105
4.5.5	Evaluasi SVM dengan SMOTE.....	107
4.5.6	Evaluasi SVM dengan Random Undersampling (RUS).....	110
4.5.7	Confusion Matrix dan tabel perbandingan performa model.....	113

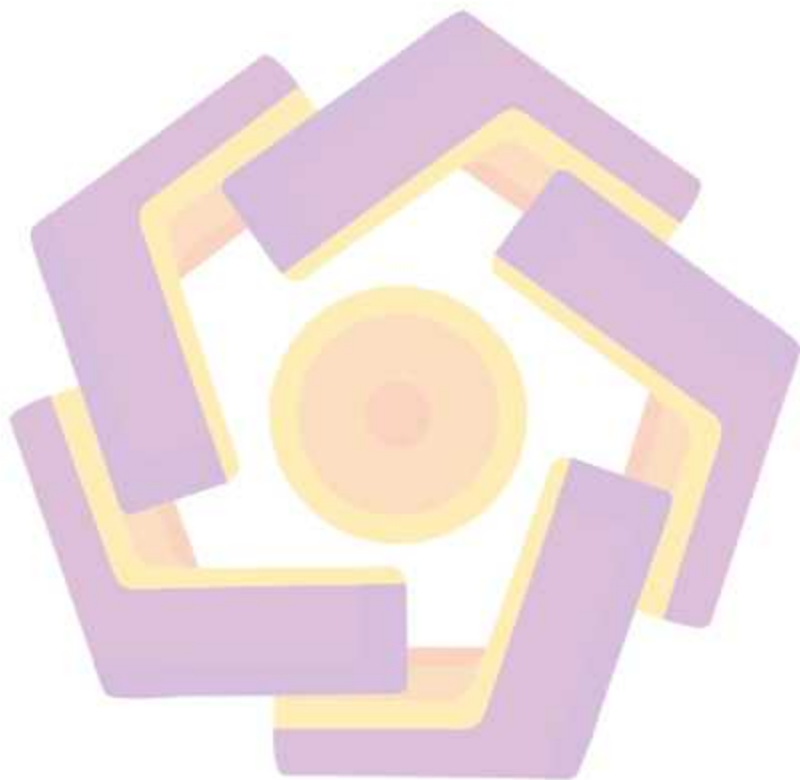
4.5.8 Analisis Precision, Recall, F1-Score, dan AUC.....	116
4.5.9 Analisis waktu komputasi & efisiensi masing-masing model	119
4.6 Analisis Temporal Sentimen	123
4.6.1 Tren bulanan sentimen positif dan negatif.....	123
4.6.2 Analisis puncak sentimen negatif pada momen politik (demo, pengesahan, revisi UU).....	125
4.6.3 Narasi longitudinal tren sentimen 2020–2023	128
4.7 Pembahasan dan Interpretasi Temuan.....	131
4.7.1 Analisis temuan utama model dan relevansi dengan data sosial..	131
4.7.2 Perbandingan hasil dengan penelitian terdahulu.....	132
BAB V PENUTUP	136
5.1 Kesimpulan.....	136
5.2 Saran.....	139
DAFTAR PUSTAKA	140
LAMPIRAN	146

DAFTAR TABEL

Tabel 2.1 Confusion Matrix.....	22
Tabel 2.2 Matriks literatur review dan posisi penelitian	25
Tabel 4.1 Distribusi tweet berdasarkan tahun.....	47
Tabel 4.3 Distribusi awal label sentimen.....	51
Tabel 4.4 Hasil transformasi case folding	52
Tabel 4.5 Hasil proses cleaning	53
Tabel 4.6 Hasil Normalisasi.....	55
Tabel 4.7 Hasil proses tokenisasi pada beberapa tweet.....	57
Tabel 4.8 Hasil penerapan stopword removal	59
Tabel 4.9 Hasil penerapan stemming.....	60
Tabel 4. 10 Rekapitulasi Tahapan Pra-Pemrosesan Teks.....	62
Tabel 4.11 Perbandingan perubahan.....	66
Tabel 4.12 Hasil Pelabelan Sentimen	68
Tabel 4.13 Distribusi Sentimen	70
Tabel 4.14 Konfigurasi	80
Tabel 4.15 Hasil Bobot TF-IDF pada Kata Dominan.....	81
Tabel 4.16 7 jenis smart features	84
Tabel 4.17 Hasil Ekstraksi Smart Features	85
Tabel 4.18 Kata dengan nilai Chi-Square tertinggi	88
Tabel 4.19 Hasil seleksi fitur	89
Tabel 4.20 Ringkasan Hasil Reduksi Fitur	90

Tabel 4.21 Hasil pengujian awal.....	91
Tabel 4.22 Data pada Tiap Split	96
Tabel 4.23 Hasil Evaluasi Model SVM Linear (Tanpa Resampling).....	99
Tabel 4.24 Distribusi hasil prediksi model	100
Tabel 4.25 Hasil performa RF pada tiga rasio pembagian data.....	102
Tabel 4.26 Confusion Matrix Model Random Forest.....	103
Tabel 4.27 Hasil performa MNB pada tiga konfigurasi pembagian data	105
Tabel 4.28 Confusion Matrix Model MNB (Split 80:20).....	107
Tabel 4.29 Hasil Evaluasi Model Setelah SMOTE	108
Tabel 4.30 Confusion Matrix SVM Setelah SMOTE.....	109
Tabel 4.31 Hasil Evaluasi Model dengan RUS	111
Tabel 4.32 Perbandingan SVM dengan dan tanpa RUS.....	113
Tabel 4.33 Confusion Matrix SVM dengan RUS (Split 80:20)	113
Tabel 4.34 SVM Linear	114
Tabel 4.35 Random Forest.....	114
Tabel 4.36 Multinomial Naïve Bayes.....	114
Tabel 4.37 Perbandingan Performa Model.....	115
Tabel 4.38 Precision	116
Tabel 4.39 Recall.....	117
Tabel 4.40 F1-Score.....	118
Tabel 4.41 <i>Area Under Curve</i>	119
Tabel 4.42 Hasil Pengukuran Waktu Komputasi.....	119
Tabel 4.43 Perbandingan Efisiensi Komputasi.....	122

Tabel 4.44 Tren Bulanan	124
Tabel 4.45 Analisis Naratif Longitudinal 2020–2023	130



DAFTAR GAMBAR

Gambar 2.1	Prosedur Umum Analisis Sentimen	9
Gambar 3.2	Diagram Alur Penelitian	39
Gambar 4.1	Distribusi Sentimen Tweet tentang Omnibus Law	47
Gambar 4.3	Wourecld negatif	52
Gambar 4.4	Wordcloud Positif	54
Gambar 4.5	Perbandingan Sebelum Dan Sesudah Normalisasi	56
Gambar 4.6	Komposisi Lexicon Sentimen	70
Gambar 4.7	Confusion Matrix (Grid 3x3 untuk semua kondisi)	95
Gambar 4.8	Perbandingan Akurasi Model per Kondisi Resampling	108
Gambar 4.9	Perbandingan Metrik Evaluasi (Accuracy, Precision, Recall, F1-Score)	110
Gambar 4.10	ROC Curves untuk Semua Model dan Kondisi	111
Gambar 4.11	Heatmap Perbandingan Akurasi	111
Gambar 4.12	Perbandingan <i>Training Time</i>	120

INTISARI

Omnibus law atau undang-undang cipta kerja telah menjadi topik perdebatan publik yang intens di Indonesia sejak disahkan pada tahun 2020. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap omnibus law melalui data Twitter dengan menggunakan pendekatan machine learning. Dataset penelitian terdiri dari 20.000 tweet mentah yang dikumpulkan dari periode 20 Juli 2020 hingga 14 April 2023, menghasilkan 17.184 tweet valid setelah penyaringan. Data melalui preprocessing intensif meliputi case folding, cleaning, normalisasi, tokenisasi, stopword removal, dan stemming, dengan reduksi 38,5% menjadi 81.705 kata unik. Pelabelan sentimen menggunakan lexicon-based menghasilkan distribusi 60,02% negatif dan 39,98% positif.

Penelitian mengimplementasikan tiga algoritma: support vector machine (svm), random forest, dan multinomial naïve bayes, dengan teknik resampling rus, ros, dan smote. Ekstraksi fitur menggunakan tf-idf dengan feature selection mereduksi dimensi dari 18.247 menjadi 7.500 fitur (58%). Hasil menunjukkan svm dengan rus menghasilkan performa terbaik: akurasi 97,65%, precision 97,59%, recall 97,45%, f1-score 97,52%, dan auc-roc 0,9929. Random forest mencapai 96,17% dan multinomial naïve bayes 94,76%. Analisis temporal mengungkapkan sentimen negatif puncak 78% pada Oktober 2020 dengan kata kunci "tolak", "buruh", "demo", dan "phk". Penelitian memberikan kontribusi metodologis berupa kerangka kerja analisis sentimen kebijakan publik berbahasa Indonesia dan secara praktis menyediakan alat monitoring opini publik real-time.

Kata kunci: analisis sentimen, Omnibus Law, Twitter, machine learning, SVM, TF-IDF, resampling

ABSTRACT

The Omnibus Law or Job Creation Law has become a topic of intense public debate in Indonesia since its enactment in 2020. This research aims to analyze Indonesian public sentiment toward the Omnibus Law through Twitter data using a machine learning approach. The research dataset consists of 20,000 raw tweets collected from July 20, 2020 to April 14, 2023, resulting in 17,184 valid tweets after filtering. The data underwent intensive preprocessing including case folding, cleaning, normalization, tokenization, stopword removal, and stemming, with a 38.5% reduction to 81,705 unique words. Lexicon-based sentiment labeling resulted in 60.02% negative and 39.98% positive distribution.

The research implements three algorithms: Support Vector Machine (SVM), Random Forest, and Multinomial Naïve Bayes, with RUS, ROS, and SMOTE resampling techniques. Feature extraction using TF-IDF with feature selection reduced dimensions from 18,247 to 7,500 features (58%). Results show SVM with RUS achieves the best performance: 97.65% accuracy, 97.59% precision, 97.45% recall, 97.52% F1-score, and 0.9929 AUC-ROC. Random Forest achieves 96.17% and Multinomial Naïve Bayes 94.76%. Temporal analysis reveals negative sentiment peaked at 78% in October 2020 with keywords "tolak/reject", "buruh/labor", "demo/protest", and "PHK/layoffs". This research provides methodological contributions in the form of a comprehensive framework for sentiment analysis of Indonesian public policy and practically provides real-time public opinion monitoring tools.

Keywords: sentiment analysis, Omnibus Law, Twitter, machine learning, SVM, TF-IDF, resampling

BAB II

PENDAHULUAN

2.1 Latar Belakang Masalah

Omnibus Law, atau Undang-Undang Cipta Kerja, telah menjadi topik yang hangat dalam diskusi publik di Indonesia sejak disahkan pada tahun 2020. Undang-undang ini bertujuan untuk menyederhanakan perizinan, meningkatkan investasi, dan merampingkan ketentuan ketenagakerjaan. Namun, respons masyarakat terhadap Omnibus Law sangat bervariasi: sebagian mendukungnya sebagai langkah maju dalam pembangunan ekonomi dan investasi, sementara lainnya menentangnya karena khawatir akan dampak negatif terhadap hak-hak buruh (Pane dkk., 2021a).

Platform media sosial seperti Twitter telah memainkan peran penting dalam memfasilitasi ekspresi dan penyebaran pandangan masyarakat tentang Omnibus Law. Dengan Twitter, pengguna dapat dengan cepat dan luas mengungkapkan sentimen mereka terhadap isu-isu kontroversial seperti ini.

Analisis sentimen menjadi alat penting dalam memahami respons publik terhadap kebijakan dan peristiwa penting (Ndruru, 2022a). Melalui analisis sentimen, tweet-tweet yang mengungkapkan pandangan positif, negatif, atau netral terhadap Omnibus Law dapat diklasifikasikan, memberikan wawasan yang berharga bagi pembuat kebijakan untuk mengevaluasi penerimaan dan dampak kebijakan tersebut di Masyarakat (Sandryan dkk., 2021a).

Masalah yang muncul dalam analisis sentimen terhadap Omnibus Law meliputi variasi dalam persepsi dan reaksi publik. Beberapa pendukung melihat Omnibus Law sebagai dorongan untuk investasi dan pertumbuhan ekonomi,

sementara yang lain merasa bahwa undang-undang ini dapat mengorbankan hak-hak tenaga kerja dan kondisi sosial-ekonomi (Adnyana dkk., 2021).

Dalam upaya mengklasifikasikan sentimen ini, algoritma Support Vector Machine (SVM) dengan kernel linier telah menjadi salah satu metode yang populer dan sering digunakan. SVM memiliki kemampuan untuk menangani data dengan dimensi tinggi dan memberikan solusi optimal dengan margin maksimal, sehingga cocok digunakan dalam tugas klasifikasi seperti analisis sentimen. Namun, meskipun SVM menawarkan keakuratan yang tinggi, algoritma ini memiliki beberapa kelemahan. Salah satu kelemahan utama SVM adalah sensitivitasnya terhadap ketidakseimbangan kelas dalam dataset. Ketika jumlah data pada satu kelas jauh lebih besar dibandingkan kelas lainnya, SVM cenderung lebih akurat dalam memprediksi kelas mayoritas, sementara performa pada kelas minoritas menurun (Rabbani dkk., 2023a).

Selain itu, Random Forest dan Multinomial Naïve Bayes juga merupakan algoritma populer dalam analisis sentimen yang menawarkan keuntungan dalam hal interpretabilitas dan efisiensi. Random Forest, dengan pendekatan ensemble-nya, mampu menangani dataset yang tidak seimbang dengan lebih baik, namun sering kali memerlukan banyak sumber daya komputasi. Di sisi lain, Multinomial Naïve Bayes dikenal karena kecepatannya dalam memproses data teks, namun asumsi independensi antar fitur yang diadopsinya dapat menjadi kelemahan dalam menangani data dengan korelasi antar fitur yang kuat (Sarimole & Ihsan, 2023).

Penelitian ini berfokus pada penerapan algoritma SVM dengan kernel linier untuk mengklasifikasikan tweet terkait Omnibus Law, dengan memperhatikan

kelemahan-kelemahan yang disebutkan di atas. Untuk meningkatkan performa klasifikasi, akan diterapkan pendekatan feature selection by correlation dan teknik resampling dengan SMOTE dan Undersampling untuk mengatasi ketidakseimbangan data. Selain itu, penelitian ini juga akan mengeksplorasi dan membandingkan performa algoritma SVM, Random Forest dan Multinomial Naïve Bayes, serta mengevaluasi efektivitas masing-masing dalam konteks analisis sentimen terhadap Omnibus Law di Twitter.

Penggunaan Natural Language Processing (NLP) akan menjadi aspek kunci dalam otomatisasi proses klasifikasi tweet berdasarkan sentimen terhadap Omnibus Law. Teknik-teknik NLP akan membantu dalam mengidentifikasi pola bahasa dan konteks yang mengungkapkan sentimen dalam teks yang diposting di Twitter. Dengan demikian, penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang sikap dan pandangan masyarakat Indonesia terhadap Omnibus Law serta berkontribusi pada pengembangan metodologi analisis sentimen yang lebih canggih dan efektif untuk kebijakan publik di masa depan.

2.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana membangun model klasifikasi sentimen yang optimal menggunakan algoritma Support Vector Machine (SVM), Random Forest, dan Multinomial Naïve Bayes untuk tweet Omnibus Law?
- b. Bagaimana performa komparasi ketiga algoritma tersebut berdasarkan metrik evaluasi accuracy, precision, recall, F1-score, dan AUC-ROC?

- c. Faktor-faktor apa yang mempengaruhi akurasi model dan bagaimana optimalisasi dapat dilakukan?

2.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah yang dirumuskan sebagai berikut:

- a. Data tweet yang dianalisis terkait Omnibus Law berasal dari platform Twitter dalam bahasa Indonesia.
- b. Data dikumpulkan melalui teknik scrapping dari tanggal 20 Juli 2020 hingga 14 April 2023.
- c. Analisis sentimen dibatasi pada dua kategori: positif dan negatif, menggunakan algoritma *Support Vector Machine (SVM)* dengan kernel linier, *Random Forest* dan *Multinomial Naïve Bayes*.
- d. Pelabelan data dilakukan secara manual untuk memastikan akurasi dalam membedakan sentimen positif dan negatif.
- e. Evaluasi kinerja algoritma terfokus pada pengukuran akurasi, presisi, recall, atau F1-score.

2.4 Tujuan Penelitian

Penelitian ini memiliki maksud dan tujuan. Adapun maksud dan tujuan dari penelitian ini adalah :

- a. Menyediakan pemahaman mendalam tentang sentimen masyarakat Indonesia terhadap Omnibus Law melalui analisis tweet di Twitter.
- b. Mengklasifikasikan tweet tentang Omnibus Law dalam kategori positif dan negatif menggunakan SVM, Random Forest dan Multinomial Naïve Bayes.

- c. Membandingkan kinerja berbagai algoritma klasifikasi untuk meningkatkan akurasi analisis sentimen.

2.5 Manfaat Penelitian

Manfaat yang diharapkan dari penyusunan penelitian ini adalah sebagai berikut :

- a. Bagi ilmu pengetahuan :

Memberikan wawasan baru tentang sentimen masyarakat terhadap Omnibus Law melalui analisis data dari Twitter, serta berkontribusi pada pengembangan metodologi dalam ilmu komputasi dan sosial.

- b. Bagi peneliti :

Menambah wawasan dalam analisis sentimen dan penerapan machine learning, serta evaluasi overfitting/underfitting melalui loss value..

- c. Bagi pengguna :

Mempercepat pengambilan keputusan terkait perubahan opini publik terhadap Omnibus Law secara efektif.

BAB III

TINJAUAN PUSTAKA

3.1 Tinjauan Pustaka

Beberapa studi terkait analisis sentimen terhadap Undang-Undang Cipta Kerja (Omnibus Law) menggunakan berbagai pendekatan salah satunya adalah penelitian mengenai dengan menggunakan SVM dalam mengidentifikasi tanggapan negatif terutama terkait ketenagakerjaan dengan akurasi 91,80% (Sukma dkk., 2020a). Namun, studi ini belum mencakup optimasi parameter SVM secara mendalam.

Sebaliknya, dalam penelitian selanjutnya berhasil meningkatkan akurasi SVM hingga 95%, meskipun menghadapi kompleksitas implementasi PSO yang memerlukan pengaturan cermat (Sabrila dkk., 2022). dan pada penelitian lainnya ditemukan bahwa kernel RBF pada SVM menawarkan keakuratan tinggi dalam klasifikasi sentimen, tetapi tetap menghadapi kendala dalam menangani ketidakseimbangan kelas dalam dataset (Rabbani dkk., 2023a). Penelitian lain juga mencatat SVM dengan kernel linier mencapai akurasi 97,05%, namun memerlukan penanganan dataset yang besar dan kompleks (Makhtum & Muhajir, 2023a).

Sementara itu, penelitian terkait sentiment dengan algoritma SVM dan Naïve Bayes, menunjukkan bahwa SVM memiliki akurasi lebih tinggi (81,31%) dibandingkan dengan Naïve Bayes (75,43%), dengan catatan perlunya pengelolaan parameter yang tepat untuk optimalisasi performa (Sarimole & Ihsan, 2023). Terakhir, Eksplorasi penggunaan PSO untuk optimalisasi parameter SVM dalam

analisis sentimen Omnibus Law di Twitter, menyoroti keefektifan PSO namun juga kompleksitas implementasi yang perlu diperhatikan dengan cermat (Augustia dkk., 2021).

Penelitian lain dalam Academic Platform Journal of Engineering and Science menunjukkan bahwa SVM memiliki akurasi yang cukup baik dalam analisis sentimen teks Twitter, yaitu sebesar 89%. Penelitian ini juga membandingkan performa SVM dengan algoritma lain seperti Random Forest dan Gaussian Naive Bayes, di mana Random Forest menghasilkan akurasi 88% dan Gaussian Naive Bayes 72%. Meskipun hasil ini menunjukkan bahwa SVM memiliki performa yang baik, penelitian ini mencatat bahwa performa pada tweet dengan sentimen negatif lebih rendah dibandingkan dengan sentimen positif dan netral, yang menunjukkan perlunya lebih banyak data untuk pelatihan yang lebih akurat.

Selain itu, penelitian yang ditemukan pada International Journal of Creative Research Thoughts menunjukkan bahwa Random Forest dapat mencapai akurasi yang lebih tinggi dibandingkan SVM dalam analisis sentimen ulasan produk di Amazon, dengan akurasi masing-masing 96,18% dan 92,33%. Studi ini menyoroti pentingnya penanganan dataset yang besar dan kompleks serta pengembangan algoritma yang lebih lanjut untuk menangani tantangan ini. Meskipun demikian, penelitian ini memperkuat relevansi penggunaan SVM dan Random Forest dalam konteks analisis sentimen, terutama dalam skenario yang melibatkan dataset yang lebih besar dan bervariasi (Akanksha dkk., 2023), (Gupta dkk., 2019).

3.2 Landasan Teori

3.2.1 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) adalah bidang multidisipliner yang memungkinkan komputer untuk memahami, menginterpretasikan, dan merespons bahasa manusia dengan cara yang bermakna dan berguna. NLP menggabungkan elemen dari berbagai disiplin ilmu seperti linguistik manusia, linguistik komputasional, rekayasa statistik, dan kecerdasan buatan dalam pembelajaran mesin. Linguistik manusia fokus pada struktur dan makna bahasa, termasuk tata bahasa dan semantik. Linguistik komputasional melibatkan pembuatan algoritma yang dapat memproses dan memahami bahasa manusia (Lee, 2024).

Rekayasa statistik menggunakan metode statistik untuk menganalisis data bahasa dan memodelkan pola bahasa. Kecerdasan buatan dalam pembelajaran mesin melibatkan pelatihan model dengan dataset besar untuk mengenali dan memprediksi pola bahasa, memungkinkan pengembangan sistem cerdas yang mampu berinteraksi dengan manusia dalam bahasa alami (Lee, 2024). Dengan memadukan berbagai disiplin ilmu ini, NLP memungkinkan pengembangan sistem yang dapat menangkap sentimen pengguna, seperti chatbot yang tidak hanya merespons perintah dasar tetapi juga dapat berinteraksi secara lebih alami dan memahami konteks percakapan, sehingga menciptakan interaksi yang lebih mendalam dan bermakna antara manusia dan komputer.

3.2.2 Analisis Sentimen

Analisis sentimen adalah metode yang digunakan untuk mengidentifikasi dan mengkategorikan opini atau perasaan yang diekspresikan dalam teks, seperti

tweet, ulasan produk, atau posting blog. Tujuan utama dari analisis sentimen adalah untuk menentukan apakah teks mengandung sentimen positif, negatif, atau netral. Proses ini melibatkan berbagai teknik pemrosesan bahasa alami (NLP) dan pembelajaran mesin. NLP digunakan untuk memecah teks menjadi unit-unit kecil, seperti kata atau frasa (tokenisasi), mengubah kata-kata ke bentuk dasarnya (stemming dan lemmatization), serta menghapus kata-kata umum yang tidak signifikan (stopword removal) (Wankhade dkk., 2022). Alur analisis sentimen dapat dilihat pada gambar 2.1.



Gambar 2.1 Prosedur Umum Analisis Sentimen

Algoritma pembelajaran mesin, seperti *Support Vector Machine* (SVM), *Random Forest*, dan *Naïve Bayes*, kemudian digunakan untuk mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya. Tantangan utama dalam analisis sentimen termasuk penanganan ironi, sarkasme, kesalahan tata bahasa, dan variasi gaya penulisan informal yang sering ditemukan dalam teks dari media sosial. Analisis sentimen juga dapat menggunakan pendekatan multimodal yang menggabungkan data dari berbagai sumber, seperti teks, suara, dan ekspresi wajah, untuk meningkatkan akurasi deteksi sentimen.

3.2.3 Twitter

Twitter adalah platform media sosial yang memungkinkan pengguna untuk mengirim dan membaca pesan berbasis teks yang dikenal sebagai "tweet." Setiap tweet dibatasi hingga 280 karakter, memungkinkan komunikasi singkat dan cepat. Twitter digunakan secara luas untuk berbagi berita, pendapat, dan informasi secara real-time, dan telah menjadi alat penting dalam berbagai bidang termasuk bisnis, politik, dan hiburan.

Twitter memiliki lebih dari 300 juta pengguna aktif bulanan, menjadikannya salah satu layanan microblogging paling populer di dunia (Wang dkk., 2022). Platform ini memungkinkan pengguna untuk mengikuti akun lain, menggunakan tagar (hashtag) untuk mengkategorikan pesan, dan terlibat dalam diskusi global tentang berbagai topik. Dalam konteks analisis sentimen, Twitter memegang peran yang sangat unik berkat kapasitasnya untuk menyediakan aliran data opini publik yang melimpah dan multi-dimensional secara waktu nyata (real-time). Karakteristik ini menjadikannya platform yang ideal bagi para akademisi dan pelaku bisnis. Mereka secara sistematis memanfaatkan data dari Twitter untuk mengidentifikasi dan menginterpretasi fluktuasi tren sentimen yang terkait dengan berbagai isu sosial, respons pasar terhadap suatu produk, maupun dampak dari suatu kejadian. Oleh karena itu, Twitter tidak hanya berfungsi sebagai medium komunikasi, tetapi juga sebagai repositori data yang sangat berharga untuk investigasi sosial dan implementasi strategi komersial (Wang dkk., 2022).

3.2.4 Emosi

Emosi merupakan respons psikologis kompleks yang melibatkan komponen kognitif, afektif, dan behavioral yang diekspresikan melalui penggunaan bahasa dalam komunikasi digital seperti media sosial Twitter (Akbar & Sugiharto, 2023). Dalam konteks analisis sentimen berbasis teks, emosi direpresentasikan melalui pemilihan kata, struktur kalimat, penggunaan tanda baca, dan elemen linguistik lainnya yang mencerminkan keadaan perasaan atau sikap pengguna terhadap suatu objek, isu, atau peristiwa. Dalam praktik analisis sentimen komputasional, spektrum emosi yang kompleks disederhanakan menjadi kategori polaritas diskrit (Ferarizki dkk., 2023). Kategori utama meliputi sentimen positif yang mencakup emosi seperti kegembiraan, persetujuan, dukungan, dan optimisme; sentimen negatif yang mencakup emosi penolakan, kritik, kemarahan, dan kekhawatiran; serta sentimen netral yang merepresentasikan pernyataan faktual tanpa muatan emosional yang jelas. Penelitian ini secara spesifik fokus pada klasifikasi biner antara sentimen positif dan negatif untuk mengidentifikasi dan mengukur polaritas opini publik terhadap kebijakan Omnibus Law di platform Twitter.

3.2.5 *Pre-Processing*

Pre-processing merupakan tahapan fundamental dan krusial dalam pipeline analisis teks yang bertujuan untuk membersihkan, menstandarisasi, dan mentransformasi data tekstual mentah menjadi format terstruktur yang dapat diproses secara efektif oleh algoritma machine learning (Khairunnisa dkk., 2021a). Kualitas hasil *pre-processing* secara langsung mempengaruhi performa model klasifikasi, sehingga setiap tahapan harus dilakukan dengan cermat dan sistematis.

Penelitian ini menerapkan lima tahapan pre-processing standar yang telah terbukti efektif dalam analisis sentimen berbahasa Indonesia.

a) *Case Folding*

Case folding adalah proses konversi seluruh karakter dalam teks menjadi huruf kecil (lowercase) untuk mencapai standarisasi representasi kata. Proses ini penting karena dalam pemrosesan teks, kata "OmnibusLaw", "omnibuslaw", dan "OMNIBUSLAW" akan diperlakukan sebagai tiga entitas berbeda jika tidak dikonversi, padahal ketiganya merujuk pada konsep yang sama. Dengan case folding, inkonsistensi kapitalisasi yang umum terjadi dalam komunikasi media sosial dapat dieliminasi, sehingga mengurangi dimensi fitur dan meningkatkan akurasi model.

b) *Cleaning*

Cleaning merupakan proses penghapusan elemen-elemen non-informatif dan noise dari teks menggunakan teknik regular expression. Elemen yang dihapus meliputi URL atau tautan web yang tidak memberikan informasi sentimen, mention (@username) yang merupakan identifikasi pengguna lain, simbol hashtag (#) sambil mempertahankan konten kata setelahnya, angka yang umumnya tidak relevan untuk klasifikasi sentimen, karakter khusus seperti tanda baca berlebihan, simbol mata uang, dan karakter non-alfabet lainnya, serta emoji dan emotikon yang meskipun berpotensi mengandung informasi emosi namun tidak digunakan dalam penelitian ini. Proses cleaning ini memastikan bahwa hanya konten tekstual substantif yang dipertahankan untuk tahap pemrosesan selanjutnya.

c) *Tokenizing*

Tokenizing atau tokenisasi adalah proses pemecahan string teks atau kalimat menjadi unit-unit kata individual yang disebut token. Proses ini mengubah representasi teks dari format string kontinyu menjadi struktur data list atau array yang berisi kata-kata terpisah. Sebagai contoh, kalimat "tolak omnibus law merugikan pekerja" akan dipecah menjadi list ["tolak", "omnibus", "law", "merugikan", "pekerja"]. Tokenisasi memungkinkan setiap kata untuk diproses, dianalisis, dan diberi bobot secara independen oleh algoritma machine learning, yang merupakan fondasi dari pendekatan bag-of-words dalam analisis teks.

d) *Stopword Removal*

Stopword removal adalah proses penghapusan kata-kata umum yang memiliki frekuensi kemunculan sangat tinggi namun tidak memberikan informasi diskriminatif signifikan untuk klasifikasi sentimen. Penelitian (Mutmainah dkk., 2024) menunjukkan bahwa penerapan stopwords removal dapat meningkatkan akurasi klasifikasi hingga 3-5% dengan mengurangi noise dan dimensi fitur. Kata-kata seperti "yang", "di", "ke", "dari", "pada", "dengan" umumnya dihapus karena bersifat fungsional gramatikal dan tidak membawa muatan sentimen. Namun, dalam penelitian ini diterapkan pendekatan selektif dengan mempertahankan kata-kata tertentu yang penting untuk analisis sentimen seperti negasi ("tidak", "bukan") dan intensifier ("sangat", "sekali", "banget") karena kata-kata tersebut berfungsi sebagai modifier yang dapat mengubah atau memperkuat polaritas sentimen kata-kata di sekitarnya.

e) *Stemming*

Stemming adalah proses reduksi morfologis yang mengkonversi kata-kata berimbuhan menjadi bentuk dasar atau akar kata (*root word*). Penelitian

(Mustikasari dkk., 2021) melakukan studi komparatif terhadap efektivitas berbagai algoritma stemming untuk dokumen berbahasa Indonesia dan merekomendasikan Sastrawi Stemmer sebagai yang paling efektif karena akurasi tinggi dan kompatibilitas dengan morfologi bahasa Indonesia. Proses stemming mengurangi variasi morfologis kata sehingga kata-kata seperti "merugikan", "dirugikan", "kerugian" semua dikonversi ke bentuk dasar "rugi". Demikian pula "pekerja", "bekerja", "pekerjaan" dikonversi ke "kerja". Reduksi ini sangat penting untuk mengurangi *sparsity* dalam matriks fitur dan memastikan bahwa kata-kata dengan akar yang sama diperlakukan sebagai satu entitas oleh model, sehingga meningkatkan generalisasi dan efisiensi komputasi.

3.2.6 Feature Extraction

Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode statistik untuk pembobotan kata yang mengukur tingkat pentingnya atau relevansi suatu term (kata) dalam sebuah dokumen relatif terhadap keseluruhan koleksi dokumen atau korpus (Mutmainah dkk., 2024). Metode ini mengatasi kelemahan pendekatan *frequency-based* sederhana dengan mempertimbangkan tidak hanya seberapa sering suatu kata muncul dalam dokumen tertentu, tetapi juga seberapa unik kata tersebut dalam konteks seluruh korpus.

Formula TF-IDF terdiri dari dua komponen utama:

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \quad (2.1)$$

Term Frequency (tf) mengukur frekuensi relatif kemunculan term t dalam dokumen d dengan membagi jumlah kemunculan term tersebut ($f_{t,d}$) dengan total

semua term dalam dokumen. Normalisasi ini mencegah bias terhadap dokumen yang lebih panjang.

$$idf(t, D) = \log \frac{N}{|\{d \in D: t \in d\}|} \quad (2.2)$$

Inverse Document Frequency (idf) mengukur tingkat keunikan atau spesifisitas suatu term dengan menghitung logaritma dari total jumlah dokumen N dibagi dengan jumlah dokumen yang mengandung term t , yaitu $|\{d \in D: t \in d\}|$. Penggunaan logaritma memberikan skala yang lebih seimbang dan menghindari pembobotan yang ekstrem.

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D) \quad (2.3)$$

Skor TF-IDF akhir merupakan produk dari kedua komponen tersebut. Kata yang memiliki frekuensi tinggi dalam satu dokumen spesifik namun jarang muncul di dokumen lain akan mendapat bobot TF-IDF tinggi, mengindikasikan bahwa kata tersebut informatif dan diskriminatif untuk dokumen tersebut. Sebaliknya, kata-kata umum yang muncul di hampir semua dokumen akan mendapat bobot mendekati nol, karena kata-kata tersebut tidak memberikan informasi yang membedakan satu dokumen dari dokumen lainnya.

Dalam konteks analisis sentimen tweet tentang Omnibus Law, kata seperti "tolak" atau "investasi" yang muncul spesifik dalam subset tweet tertentu akan mendapat bobot tinggi, sementara kata umum seperti "dan" atau "yang" mendapat bobot rendah karena muncul di hampir semua dokumen.

3.2.7 *Lexicon Based Sentiment Analysis*

Pendekatan analisis sentimen berbasis leksikon (*lexicon-based approach*) merupakan metode yang memanfaatkan kamus atau dictionary kata-kata yang telah

diberi label atau skor sentimen untuk mengidentifikasi dan mengkuantifikasi polaritas emosional dalam teks (Firdaus & Herdiani, 2024). Berbeda dengan pendekatan supervised machine learning yang memerlukan data training berlabel dalam jumlah besar, metode lexicon-based dapat diterapkan tanpa fase training dan lebih fleksibel untuk disesuaikan dengan konteks domain spesifik dan karakteristik linguistik bahasa Indonesia.

Proses scoring sentimen dilakukan dengan menghitung skor positif dan skor negatif secara terpisah. Setiap dokumen akan memiliki dua skor:

$$S = \sum_{i=1}^n w_i \quad (2.4)$$

$$\begin{aligned} S_{positive} &= S > 0 \\ S_{negative} &= S < 0 \\ S_{neutral} &= S = 0 \end{aligned} \quad (2.5)$$

Proses pelabelan dilakukan menggunakan pendekatan lexicon based, di mana hasil klasifikasi dianalisis untuk mengidentifikasi jenis opini yang paling dominan. Setiap kalimat dianalisis untuk mengekstraksi nilai sentimen, kemudian kata-kata dalam teks diklasifikasikan menggunakan persamaan Formula ini memungkinkan klasifikasi yang lebih nuanced dengan membandingkan kekuatan relatif antara kata-kata positif dan negatif dalam dokumen. Pendekatan ini lebih robust dibandingkan simple threshold karena mempertimbangkan balance antara kedua polaritas.

3.2.8 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah metode yang banyak digunakan dan berpengaruh dalam menyelesaikan masalah klasifikasi. Metode ini pertama kali diperkenalkan pada tahun 1992 di Workshop on Computational Learning Theory dan didasarkan pada teori komputasi seperti margin hyperplane. SVM adalah metode klasifikasi berbasis linear yang menyelesaikan masalah menggunakan persamaan Lagrangian, bentuk ganda dari SVM melalui pemrograman kuadrat. SVM dianggap sebagai alternatif yang efisien dalam memproses data skala besar. Algoritma SVM mampu membedakan antara kelas-kelas yang berbeda dengan efektif menggunakan konsep margin (Makhtum & Muhajir, 2023b). Margin mewakili jarak antara dua kelas data terdekat ke garis hyperplane, sementara hyperplane adalah pemisah antara kelas-kelas data yang berbeda. Hyperplane dengan margin terbaik menyediakan generalisasi yang berguna untuk hasil klasifikasi yang lebih baik. Algoritma SVM mencari hyperplane yang memaksimalkan jarak margin, yang dikenal sebagai Maximum Marginal Hyperplane (MMH).

SVM adalah metode klasifikasi yang membagi data menjadi dua kelas yang berbeda. Formula untuk klasifikasi data dengan SVM menggunakan notasi y_i untuk mewakili kelas data ke- i :

$$w \cdot x_i + b \geq +1, \text{ untuk } y_i \quad (2.6)$$

$$w \cdot x_i + b \leq -1, \text{ untuk } y_i \quad (2.7)$$

Dalam formula tersebut, w adalah vektor normal hyperplane, dan b adalah posisi hyperplane relatif terhadap pusat koordinat. Jarak hyperplane dari titik pusat ke titik data adalah:

$$\min_z; w = \frac{1}{2} |w|^2 \rightarrow \min ; \quad (2.8)$$

$$y_i(w \cdot x_i + b) \geq 1 \quad (2.9)$$

Dengan kata lain, SVM mencoba untuk memisahkan data kelas positif dan negatif dengan margin yang maksimal, sehingga meningkatkan kemampuan generalisasi model dalam klasifikasi data baru.

3.2.9 *Random Forest*

Random Forest (RF) pertama kali diperkenalkan oleh Leo Breiman pada tahun 2001. RF adalah metode pembelajaran mesin yang menggunakan ensemble learning untuk meningkatkan akurasi prediksi dengan membangun banyak pohon keputusan (*decision tree*) secara acak. Setiap pohon keputusan dalam hutan acak dikembangkan dari subset data yang berbeda, dan hasil akhir diperoleh dengan cara menggabungkan prediksi dari semua pohon keputusan dalam ensemble tersebut (Azhari dkk., 2021a). *Random Forest* terdiri dari sekumpulan pohon keputusan yang digunakan untuk mengklasifikasi data ke dalam suatu kelas. Pohon keputusan dibangun dengan menentukan node akar dan berakhir dengan beberapa node daun untuk mendapatkan hasil akhir. Proses pembentukan pohon keputusan dalam metode RF mirip dengan proses pada *Classification and Regression Tree* (CART), namun pada RF tidak dilakukan pemangkasan (*pruning*).

Indeks Gini digunakan untuk memilih fitur di setiap simpul internal dari pohon keputusan. Indeks Gini mengukur ketidakmurnian atau heterogenitas sebuah simpul. Nilai Indeks Gini dapat dihitung sebagai berikut:

$$Gini(D) = 1 - \sum_{i=1}^n p_i^2 \quad (2.10)$$

Di mana p_i adalah proporsi sampel yang termasuk dalam kelas i di simpul tersebut. Semakin kecil nilai Indeks Gini, semakin murni simpul tersebut.

3.2.10 *Multinomial Naïve Bayes*

Multinomial Naïve Bayes adalah salah satu algoritma pembelajaran mesin berbasis probabilistik yang sering digunakan untuk klasifikasi teks, terutama ketika fitur-fitur yang digunakan adalah kata-kata atau n-gram. Algoritma ini mengasumsikan bahwa setiap fitur (misalnya, kata dalam teks) bersifat independen satu sama lain, yang dikenal sebagai asumsi kemandirian bersyarat (Sabrani dkk., 2020). *Multinomial Naïve Bayes* menghitung probabilitas dari setiap kelas berdasarkan frekuensi kemunculan fitur dalam data latih. Rumus dasar dari algoritma ini adalah:

$$P(c | d) = \frac{P(c) \cdot P(d | c)}{P(d)} \quad (2.10)$$

Dengan menggunakan *Multinomial Naïve Bayes*, kita dapat menghitung probabilitas dari setiap kelas untuk dokumen baru dan mengklasifikasikannya ke kelas dengan probabilitas tertinggi.

3.2.11 *SMOTE (Synthetic Minority Over-sampling Technique)*

SMOTE (Synthetic Minority Over-sampling Technique) adalah salah satu teknik resampling yang paling efektif dan banyak digunakan untuk mengatasi

masalah ketidakseimbangan kelas dalam dataset. Berbeda dengan metode oversampling sederhana yang hanya menduplikasi sampel yang ada, SMOTE mengambil pendekatan yang lebih cerdas dengan menciptakan sampel baru secara sintetis di kelas minoritas. Cara kerjanya adalah dengan memilih satu sampel dari kelas minoritas dan mengidentifikasi tetangga-tetangganya yang terdekat. Kemudian, SMOTE akan membuat sampel baru di antara mereka dengan menggunakan interpolasi—secara efektif menciptakan titik data baru di sepanjang garis yang menghubungkan sampel asli dengan tetangganya. Teknik ini tidak hanya meningkatkan jumlah data di kelas minoritas, tetapi juga memperkaya variasi dan informasi yang dimiliki oleh kelas tersebut, yang pada akhirnya membantu model pembelajaran mesin untuk mempelajari pola kelas minoritas dengan lebih baik.

Dengan menciptakan data minoritas yang beragam, SMOTE secara signifikan mengurangi risiko overfitting yang sering terjadi pada teknik oversampling yang hanya menggandakan data. Overfitting terjadi ketika model terlalu berfokus pada sampel yang berulang sehingga gagal menggeneralisasi dengan baik pada data baru. Karena SMOTE menciptakan sampel baru yang unik, model dapat belajar dari variasi yang lebih luas tanpa harus menghafal sampel yang sama.

3.2.12 *Resampling dengan Penyeimbangan Ulang (Under-sampling dan Oversampling)*

Dalam ranah pembelajaran mesin, dataset yang tidak seimbang (imbalanced dataset) sering menjadi tantangan signifikan yang dapat memengaruhi kinerja model. Untuk mengatasi bias ini, teknik penyeimbangan ulang (resampling)

diterapkan. Metode ini secara fundamental bertujuan untuk memodifikasi distribusi kelas dalam dataset sehingga model dapat belajar dari representasi yang lebih adil dari setiap kategori. Ada dua pendekatan utama yang digunakan dalam proses ini: under-sampling dan over-sampling.

Under-sampling adalah strategi yang fokus pada pengurangan jumlah sampel dari kelas mayoritas. Pendekatan ini secara efektif mengurangi redundansi data dan mempercepat proses pelatihan model. Meskipun berpotensi mengurangi overfitting—sebuah kondisi di mana model terlalu spesifik terhadap data pelatihan—teknik ini membawa risiko inheren yaitu hilangnya informasi penting yang mungkin terkandung dalam sampel yang dibuang. Oleh karena itu, pemilihan sampel yang akan dihapus harus dilakukan dengan cermat untuk memastikan bahwa karakteristik esensial dari kelas mayoritas tetap terwakili.

Sebaliknya, over-sampling beroperasi dengan menambah jumlah sampel di kelas minoritas, sehingga ukurannya setara dengan kelas mayoritas. Metode ini sangat efektif dalam memberikan model lebih banyak data untuk mempelajari pola-pola yang unik pada kelas minoritas. Pendekatan ini dapat dilakukan secara sederhana dengan menggandakan sampel yang ada, atau secara lebih canggih menggunakan algoritma seperti SMOTE (*Synthetic Minority Over-sampling Technique*), yang menciptakan sampel baru yang sintesis berdasarkan pola-pola yang diamati dari sampel minoritas yang sudah ada. Namun, over-sampling juga memiliki kelemahan; penggandaan atau sintesis sampel dapat berpotensi menyebabkan overfitting karena model mungkin menjadi terlalu spesifik terhadap sampel yang dibuat, mengurangi kemampuannya untuk menggeneralisasi pada data

yang benar-benar baru. Dalam praktik profesional, sering kali dilakukan kombinasi dari kedua metode ini untuk mencapai keseimbangan optimal. Pendekatan hibrida ini, yang menggabungkan pengurangan data dari kelas mayoritas dengan penambahan data pada kelas minoritas, sering kali menghasilkan model yang paling akurat dan kokoh, mampu mengelola dataset tidak seimbang tanpa mengorbankan informasi atau memicu overfitting.

3.2.13 Evaluasi Model

Evaluasi model adalah langkah penting dalam siklus hidup pengembangan model pembelajaran mesin. Tujuan dari evaluasi adalah untuk mengukur kinerja model menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Confusion matrix adalah matriks kontingensi 2x2 yang merupakan tools fundamental dalam evaluasi performa model klasifikasi biner dengan cara menampilkan distribusi hasil prediksi model dibandingkan dengan label aktual, sehingga memberikan gambaran komprehensif tentang jenis-jenis kesalahan yang dibuat oleh model dan menjadi dasar untuk perhitungan berbagai metrik evaluasi (Pratama & Marjun, 2022).

Struktur confusion matrix:

Tabel 2.1 Confusion Matrix

Prediksi Positif	Prediksi Negatif	
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

Empat sel dalam matriks merepresentasikan:

- *True Positive* (TP): Jumlah instance positif yang diprediksi dengan benar sebagai positif
- *True Negative* (TN): Jumlah instance negatif yang diprediksi dengan benar sebagai negative
- *False Positive* (FP): Jumlah instance negatif yang salah diprediksi sebagai positif (Type I Error)
- *False Negative* (FN): Jumlah instance positif yang salah diprediksi sebagai negatif (Type II Error)

Dari confusion matrix, dapat diturunkan berbagai metrik evaluasi:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2.10)$$

Accuracy mengukur proporsi prediksi yang benar dari total seluruh prediksi. Metrik ini memberikan gambaran overall performance namun dapat misleading pada dataset tidak seimbang.

$$Precision = TP / (TP + FP) \quad (2.11)$$

Precision mengukur ketepatan prediksi positif, menjawab pertanyaan "dari semua instance yang diprediksi positif, berapa persen yang benar-benar positif?" Metrik ini penting ketika cost of false positive tinggi.

$$Recall = TP / (TP + FN) \quad (2.12)$$

Recall (juga disebut *sensitivity* atau *true positive rate*) mengukur kemampuan model menangkap instance positif, menjawab pertanyaan "dari semua instance positif aktual, berapa persen yang berhasil terdeteksi?" Metrik ini krusial ketika cost of false negative tinggi.

$$F1 - Score = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (2.13)$$

F1-Score adalah harmonic mean dari precision dan recall yang memberikan satu nilai metrik yang menyeimbangkan keduanya. Penelitian (Pratama & Marjun, 2022) menjelaskan bahwa F1-Score sangat berguna untuk dataset tidak seimbang karena memberikan penalti pada model yang memiliki precision atau recall yang sangat rendah meskipun accuracy-nya tinggi.

Model kemudian dilatih dan diuji berulang kali, menggunakan fold yang berbeda untuk pelatihan dan pengujian pada setiap iterasi. Proses ini memberikan gambaran yang lebih akurat tentang kinerja model, karena hasilnya merupakan rata-rata dari berbagai pengujian.



3.3 Keaslian Penelitian

Tabel 2.2 Matriks literatur review dan posisi penelitian

Analisis Sentimen Tentang Omnibus Law di Twitter

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Analisis Sentimen Omnibus Law di Twitter Menggunakan SVM dengan Kernel Linier	Eki Aidio Sukma, Journal of Social Media Studies, 2020	Klasifikasi Sentimen pada Tweet Omnibus Law dengan SVM dengan Kernel Linier	SVM efektif mengidentifikasi sentimen negatif terkait omnibus law dengan akurasi 91.80%.	Pertunya optimasi parameter SVM yang lebih mendalam.	Penelitian Eki menggunakan Support Vector Machine dengan kernel linier dan mencapai akurasi 91.80%. Penelitian yang dilakukan penulis akan melakukan optimasi parameter dan penggunaan teknik feature

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
						selection untuk meningkatkan akurasi lebih lanjut. Selain itu, akan dibandingkan performa Support Vector Machine dengan Random Forest dan Multinomial Naïve Bayes.
2	Sentiment Analysis of Omnibus Law Using SVM with Linear Kernel	Makhtum & Muhajir, BAREKENG: Jurnal Ilmu Matematika	Klasifikasi Sentimen pada Tweet Omnibus Law dengan SVM Kernel Linier	SVM kernel linier memberikan akurasi tertinggi atau 97.05% dalam klasifikasi sentimen.	Memerlukan penanganan dataset yang besar dan kompleks.	Penelitian Makhtum & Muhajir fokus pada Support Vector Machine dengan kernel linier dan mencapai akurasi 97.05% dalam klasifikasi sentimen.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		dan Terapan, 2023				<p>Penelitian yang dilakukan penulis, selain menggunakan Support Vector Machine dengan kernel linier, juga akan menerapkan teknik resampling seperti Synthetic Minority Oversampling Technique (SMOTE) dan Adaptive Boosting (AdaBoost) untuk mengatasi ketidakseimbangan data serta</p>

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
						membandingkan dengan algoritma lain.
3	Sentiment Analysis of Labor Reform in Indonesia using Naive Bayes and Support Vector Machine	Matheos Sarimole & Ihsan, Indonesian Journal of Computational Linguistics, 2023	Klasifikasi Sentimen pada Tweet Omnibus Law dengan SVM dan Naive Bayes	SVM menunjukkan akurasi lebih tinggi dibandingkan Naive Bayes.	Perlu adanya pengelolaan parameter yang tepat untuk optimalisasi performa SVM.	Penelitian Matheos Sarimole & Ihsan membandingkan Support Vector Machine dan Naive Bayes dalam konteks reformasi ketenagakerjaan. Penelitian yang dilakukan penulis, selain menggunakan Support Vector Machine dan Naive Bayes, juga akan menerapkan Random Forest

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
						untuk perbandingan performa dalam konteks Omnibus Law di Twitter.
4	Sentiment Analysis of Omnibus Law on Twitter Using Support Vector Machine	Eki Aidio Sukma, Journal of Social Media Studies, 2020	Klasifikasi Sentimen pada Tweet Omnibus Law dengan SVM	Banyak tanggapan negatif, terutama terkait dengan aspek ketenagakerjaan.	SVM kurang efektif dalam mengklasifikasikan sentimen.	Penelitian Eki Aidio Sukma fokus pada Support Vector Machine dalam analisis sentimen Omnibus Law. Penelitian yang dilakukan penulis, selain menggunakan Support Vector Machine, juga akan menerapkan teknik feature selection dan optimasi

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
						parameter untuk meningkatkan akurasi. Akan dibandingkan dengan algoritma Random Forest dan Multinomial Naïve Bayes.
5	Analisis Sentimen Terhadap Undang-Undang Cipta Kerja Menggunakan Metode Naïve Bayes	Gusti Naufhal Daffa Adnyana, I., Adams, F., & Windari Oktavia, A. (2021).	Klasifikasi Sentimen pada Tweet Omnibus Law dengan Naïve Bayes	Naïve Bayes dapat digunakan untuk mengklasifikasikan sentimen terhadap Omnibus Law dengan akurasi yang memadai.	Perlu perhatian dalam manajemen data dan parameter untuk optimalisasi performa.	Penelitian Gusti Naufhal Daffa Adnyana menggunakan Naïve Bayes untuk analisis sentimen. Penelitian yang dilakukan penulis membandingkan Naïve Bayes dengan Support Vector

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta- Indonesia.				Machine untuk menentukan algoritma terbaik dalam klasifikasi sentimen Omnibus Law.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
6	Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM	Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023). MALCOM: Indonesian Journal of Machine Learning and Computer	Klasifikasi Sentimen terhadap kenaikan harga BBM	Akuari lebih tinggi dengan Kernel RBF	Tidak seimbang dataset	Penelitian Rabbani fokus pada evaluasi kernel Support Vector Machine. Penelitian yang dilakukan penulis menggunakan Support Vector Machine dengan kernel linier dan teknik resampling untuk mengatasi ketidakseimbangan data dalam konteks analisis sentimen Omnibus Law. Selain itu, akan dibandingkan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Science, 3(2), 153–160.				dengan Random Forest dan Multinomial Naïve Bayes.
7	Sentiment Analysis of Twitter Texts Using Machine Learning Algorithms	Hawar Sameen Ali Barzenji, Academic Platform Journal of Engineering and Science, 2021	Menganalisis sentimen teks Twitter menggunakan algoritma pembelajaran mesin seperti SVM, Random Forest, dan	SVM mencapai akurasi 89%, Random Forest 88%, dan Gaussian Naive Bayes 72%.	Performa pada tweet negatif lebih rendah. Dibandingkan tweet positif dan netral. Memerlukan lebih banyak data untuk pelatihan yang lebih akurat.	Penelitian Hawar Sameen Ali Barzenji memberikan perbandingan antara Support Vector Machine, Random Forest, dan Naive Bayes. Penelitian yang dilakukan penulis juga menggunakan Support Vector Machine dan membandingkannya dengan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			Gaussian Naive Bayes.			algoritma lain seperti Random Forest dan Multinomial Naive Bayes dalam konteks analisis sentimen Omnibus Law.
8	Sentiment Analysis Of Social-Media Using Machine Learning Approach	Akanksha M. Dhole, Neeraj Sahu, International Journal of Creative Research	Mengklasifikasikan sentimen dari ulasan produk di Amazon menggunakan Random Forest dan SVM.	Random Forest mencapai akurasi 96.18%, sedangkan SVM mencapai 92.33%.	Perlu pengembangan lebih lanjut untuk menangani dataset besar dan kompleks.	Penelitian Akanksha M. Dhole & Neeraj Sahu berfokus pada analisis sentimen ulasan produk di Amazon, sementara penelitian yang dilakukan penulis fokus pada analisis sentimen terkait Omnibus Law di Twitter.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Thoughts (IJCRT), 2023				Namun, pendekatan yang digunakan untuk membandingkan algoritma Support Vector Machine dan Random Forest sangat relevan dan serupa dengan pendekatan penelitian penulis.

BAB IV

METODE PENELITIAN

4.1 Metode Pengumpulan Data

Dalam penelitian ini, digunakan metode pengumpulan data dengan teknik scraping pada platform media sosial Twitter. Tagar (#) tertentu seperti #CiptaKerja, #OmnibusLaw, #tolakomnibuslaw, #gagalkanomnibuslaw, #RUUPenindasan, dan #CabutUUCiptaKerja dimanfaatkan untuk mengumpulkan total 20.000 tweet. Proses pengumpulan data dilakukan dalam periode dari Juli 2020 hingga April 2023, memastikan cakupan waktu yang cukup untuk menangkap berbagai opini publik terkait topik ini. Fitur yang diambil dari tweet tersebut meliputi ID pengguna, teks tweet, dan tanggal tweet. Selain itu, data juga dikumpulkan berdasarkan hasil dari Google Trends untuk memastikan bahwa tagar yang digunakan relevan dan sering dicari oleh pengguna selama periode penelitian. Penggunaan Google Trends membantu mengidentifikasi tagar yang sedang populer dan banyak dibicarakan, sehingga data yang diperoleh lebih representatif terhadap opini publik yang berkembang. Dengan demikian, kombinasi teknik scraping dan analisis Google Trends memberikan dasar yang kuat untuk analisis sentimen yang akurat.

4.2 Metode Analisis Data

Metode analisis data dalam penelitian ini melibatkan beberapa tahapan mulai dari pra-pemrosesan teks hingga evaluasi model. Tahap pertama adalah pra-pemrosesan teks, yang mencakup filtering untuk menghapus tweet yang tidak relevan seperti spam atau duplikasi, case folding untuk mengubah semua huruf

dalam teks tweet menjadi huruf kecil guna memastikan konsistensi, tokenization untuk memecah teks tweet menjadi kata-kata atau token, stopword removal untuk menghapus kata-kata umum yang tidak memiliki makna signifikan seperti "dan", "di", "yang", simbol, serta stemming untuk mengubah kata-kata menjadi bentuk dasarnya. Setelah melalui tahap pra-pemrosesan, data dibagi menjadi dua set, yaitu data training dan data testing. Data training digunakan untuk melatih model, sementara data testing digunakan untuk menguji performa model.

Pada tahap selanjutnya, dilakukan pemilihan fitur menggunakan teknik correlation-based feature selection untuk memilih fitur yang paling relevan dan signifikan dalam klasifikasi sentimen. Dalam tahap klasifikasi, berbagai algoritma pembelajaran mesin diterapkan, termasuk Support Vector Machine (SVM) dengan kernel linier, Random Forest, dan Multinomial Naïve Bayes. Untuk meningkatkan performa klasifikasi, selain teknik SMOTE (*Synthetic Minority Over-sampling Technique*) yang digunakan untuk mengatasi ketidakseimbangan kelas, juga diterapkan teknik boosting lainnya seperti AdaBoost dan Gradient Boosting. Teknik-teknik ini digunakan untuk membandingkan dan meningkatkan akurasi model dengan menggabungkan kekuatan beberapa model sederhana untuk membentuk model yang lebih kuat. Evaluasi model dilakukan dengan menggunakan confusion matrix untuk menghitung metrik evaluasi seperti true positives, true negatives, false positives, dan false negatives. Dari hasil ini, metrik akurasi, presisi, recall, dan F1-score dihitung untuk menilai performa model.

4.3 Jenis, Sifat dan Pendekatan Penelitian

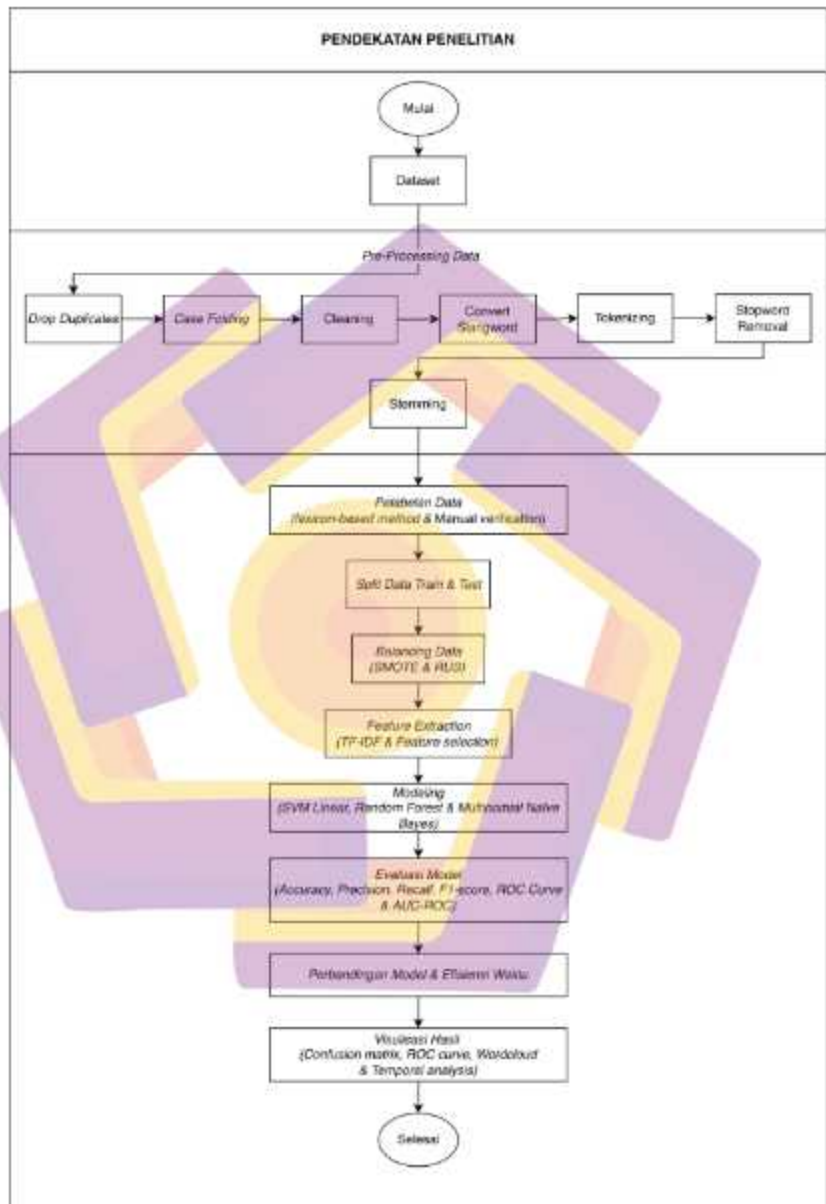
Penelitian yang diterapkan dalam studi ini merupakan penelitian Eksperimen. Penelitian eksperimen dilakukan dengan menerapkan serangkaian tindakan untuk membuktikan suatu konsep. Dalam konteks ini, penelitian ini bertujuan untuk menguji efektivitas berbagai algoritma pembelajaran mesin dalam analisis sentimen terhadap Omnibus Law di Twitter.

Sifat penelitian ini adalah deskriptif. Penelitian deskriptif bertujuan untuk menggambarkan atau menganalisis karakteristik, sifat, atau kondisi dari fenomena yang diteliti. Penelitian ini akan menggambarkan dan menganalisis hasil pengujian dari berbagai algoritma pembelajaran mesin dalam mengklasifikasikan sentimen masyarakat terhadap Omnibus Law.

Pendekatan penelitian yang digunakan adalah kuantitatif. Pendekatan kuantitatif lebih berfokus pada pengukuran dan analisis statistik untuk menguji hipotesis dan menarik generalisasi tentang fenomena yang diteliti. Data dalam penelitian ini dikumpulkan melalui teknik scraping dari platform Twitter, kemudian dianalisis menggunakan berbagai algoritma pembelajaran mesin seperti SVM, Random Forest, dan Multinomial Naïve Bayes, serta teknik resampling untuk mengatasi ketidakseimbangan data. Hasil analisis akan diukur dengan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.

4.4 Alur Penelitian

Alur penelitian yang diaplikasikan pada penelitian ini diunjukkan pada gambar sebagai berikut :



Gambar 3.2 Diagram Alur Penelitian

Berikut penjelasan pada alur penelitian pada Gambar 3.2 sebagai berikut:

Tahap awal penelitian ini melibatkan beberapa langkah penting untuk memastikan validitas dan relevansi penelitian yang akan dilakukan. Proses dimulai dengan validasi penelitian tentang permasalahan yang akan dianalisis, di mana peneliti mengidentifikasi dan memastikan bahwa topik penelitian memang layak untuk diteliti.

1. Pendekatan Penelitian (Validasi Penelitian, Studi Literatur, Identifikasi Penelitian)

Tahap awal penelitian melibatkan validasi penelitian tentang permasalahan yang akan dianalisis, studi literatur untuk mengetahui masalah yang akan dianalisis serta algoritma yang tepat berdasarkan studi literatur, dan identifikasi penelitian untuk menentukan masalah penelitian dan pertanyaan-pertanyaan yang akan dijawab.

2. Metode Pengumpulan Data (Pemilihan Data Latih dan Uji, Pengumpulan Data, Pre-Processing Data)

Pada tahap pengumpulan data, peneliti melakukan berbagai upaya untuk memperoleh data yang akurat dan dapat diandalkan. Data dikumpulkan melalui teknik scraping dari platform media sosial Twitter dengan menggunakan tagar (#) tertentu yang relevan dengan topik penelitian, seperti #CiptaKerja, #OmnibusLaw, #tolakomnibuslaw, dan lainnya. Data yang dikumpulkan meliputi ID pengguna, teks tweet, dan tanggal tweet. Setelah data dikumpulkan, dilakukan proses pre-processing untuk membersihkan dan mempersiapkan data untuk analisis lebih lanjut.

Proses pre-processing mencakup beberapa tahapan sistematis sebagai berikut:

a. Drop Duplicates

Tahap ini bertujuan untuk menghilangkan data duplikat agar tidak terjadi pengulangan data yang dapat mempengaruhi hasil analisis.

b. Case Folding

Proses mengubah seluruh huruf dalam teks menjadi huruf kecil (lowercase) agar tidak terjadi perbedaan makna akibat perbedaan penulisan huruf kapital.

c. Cleaning

Tahap ini dilakukan untuk menghapus karakter yang tidak diperlukan seperti mention (@username), hashtag (#), URL, angka, tanda baca, emoji, serta karakter khusus lainnya.

d. Convert Slangword (Normalisasi Kata Tidak Baku)

Kata-kata tidak baku atau bahasa gaul diubah menjadi kata baku menggunakan kamus slangword, misalnya "gk" menjadi "tidak", "aja" menjadi "saja", dan sejenisnya.

f. Tokenizing

Teks dipecah menjadi token atau kata-kata individu agar dapat diproses lebih lanjut oleh sistem.

g. Stopword Removal

Tahap ini bertujuan untuk menghapus kata-kata umum yang tidak memiliki makna sentimen signifikan seperti “yang”, “dan”, “di”, “ke”, “dari”, serta kata umum lainnya.

h. Stemming

Setelah stopword dihapus, dilakukan proses stemming untuk mengubah setiap kata menjadi kata dasar dengan menghilangkan imbuhan seperti awalan, akhiran, sisipan, dan konfiks.

3. Pelabelan Data (Lexicon-based Method & Manual Verification)

Setelah pre-processing selesai, dilakukan pelabelan data menggunakan kombinasi metode lexicon-based dan verifikasi manual. Pendekatan lexicon-based menggunakan kamus sentimen untuk memberikan label awal secara otomatis, kemudian dilakukan verifikasi manual oleh peneliti untuk memastikan akurasi pelabelan, terutama untuk tweet yang mengandung sarkasme, ironi, atau konteks yang kompleks.

4. Split Data Train & Test

Data yang telah dilabeli kemudian dibagi menjadi data latih (training data) dan data uji (testing data) dengan berbagai rasio pembagian yang diuji dalam penelitian ini, seperti 80:20, 85:15, dan 90:10, untuk mengevaluasi performa model pada berbagai konfigurasi.

5. Balancing Data (SMOTE & RUS)

Untuk menangani ketidakseimbangan kelas jika ada, penelitian ini menguji dua teknik balancing data yaitu Synthetic Minority Over-sampling Technique

(SMOTE) untuk oversampling kelas minoritas dan Random Undersampling (RUS) untuk undersampling kelas mayoritas.

6. Feature Extraction (TF-IDF & Feature Selection)

Tahap ekstraksi fitur menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) dengan n-gram (1,3) untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning. Selain itu, dilakukan feature selection menggunakan metode Chi-Square untuk memilih fitur-fitur yang paling relevan dan mengurangi dimensionalitas data, sehingga meningkatkan efisiensi komputasi dan performa model.

7. Modeling (SVM Linear, Random Forest & Multinomial Naive Bayes)

Pada tahap modeling, tiga algoritma klasifikasi diimplementasikan dan dibandingkan performanya, yaitu Support Vector Machine (SVM) Linear yang efektif untuk data berdimensi tinggi, Random Forest yang memberikan interpretabilitas tinggi melalui feature importance, dan Multinomial Naive Bayes yang menawarkan kecepatan pelatihan sangat tinggi sebagai baseline model.

8. Evaluasi Model (Accuracy, Precision, Recall, F1-score, ROC Curve & AUC-ROC)

Evaluasi model dilakukan secara komprehensif menggunakan berbagai metrik evaluasi untuk memastikan reliabilitas hasil. Metrik yang digunakan meliputi accuracy untuk mengukur ketepatan prediksi secara keseluruhan, precision untuk mengukur ketepatan prediksi positif, recall untuk mengukur kemampuan model menemukan semua instance positif, F1-score sebagai harmonic mean dari

precision dan recall, serta ROC Curve dan AUC-ROC untuk mengukur kemampuan diskriminasi kelas model pada berbagai threshold.

9. Perbandingan Model & Efisiensi Waktu

Dilakukan analisis perbandingan performa ketiga model pada berbagai konfigurasi untuk menentukan model terbaik. Selain akurasi, efisiensi waktu pelatihan juga menjadi pertimbangan penting untuk implementasi sistem real-time, di mana SVM Linear menunjukkan keseimbangan optimal antara akurasi tinggi (97,65%) dan waktu pelatihan yang efisien (41,58 detik).

10. Visualisasi Hasil (Confusion Matrix, ROC Curve, Wordcloud & Temporal Analysis)

Tahap akhir adalah visualisasi hasil untuk memudahkan interpretasi dan komunikasi temuan penelitian. Confusion matrix digunakan untuk menunjukkan detail prediksi benar dan salah pada setiap kelas, ROC Curve untuk memvisualisasikan trade-off antara true positive rate dan false positive rate, wordcloud untuk menampilkan kata-kata dominan pada setiap kelas sentimen, dan temporal analysis untuk menganalisis dinamika perubahan sentimen publik dari waktu ke waktu yang menunjukkan pengaruh momen politik terhadap opini publik

BAB V

HASIL DAN PEMBAHASAN

5.1 Pengumpulan dan Karakteristik Data

5.1.1 Deskripsi data

Penelitian ini menggunakan data sekunder yang diperoleh dari platform Twitter sebagai sumber utama untuk menganalisis opini publik terkait *Undang-Undang Cipta Kerja (Omnibus Law)*. Twitter dipilih karena merupakan salah satu media sosial dengan tingkat interaksi tinggi di Indonesia dan sering menjadi ruang publik digital tempat masyarakat mengekspresikan pandangannya terhadap isu sosial, ekonomi, dan politik.

Proses pengumpulan data dilakukan melalui teknik web scraping menggunakan kombinasi pustaka Tweepy dan Snsrape pada bahasa pemrograman Python. Penerapan dua alat ini memungkinkan peneliti untuk memperoleh data dengan volume besar, terstruktur, dan dapat difilter berdasarkan waktu, kata kunci, serta bahasa.

Adapun kata kunci (hashtag) yang digunakan untuk mengumpulkan tweet meliputi:

#OmnibusLaw, #CiptaKerja, #TolakOmnibusLaw, #GagalkanOmnibusLaw, #CabutUUCiptaKerja, dan #RUUPenindasan.

Pemilihan tagar tersebut dilakukan setelah peneliti menganalisis tren pencarian publik melalui Google Trends, sehingga memastikan bahwa topik yang dikumpulkan benar-benar mewakili isu sosial yang sedang ramai diperbincangkan.

Proses pengumpulan data dilakukan selama periode 20 Juli 2020 hingga 14 April 2023, dengan tujuan menangkap dinamika opini publik sejak masa awal pembahasan hingga fase penyesuaian pasca pengesahan undang-undang. Rentang waktu ini juga memungkinkan analisis tren temporal untuk melihat perubahan sentimen publik sepanjang tiga tahun kebijakan ini diberlakukan.

Dari hasil *scraping*, terkumpul sebanyak 20.000 tweet mentah, yang kemudian melalui proses pembersihan (*cleaning*) dan validasi untuk menghapus data duplikat, tweet non-Bahasa Indonesia, serta unggahan bot yang bersifat spam atau tidak relevan dengan konteks kebijakan. Setelah melalui tahap penyaringan tersebut, diperoleh 14.874 tweet valid yang digunakan dalam analisis.

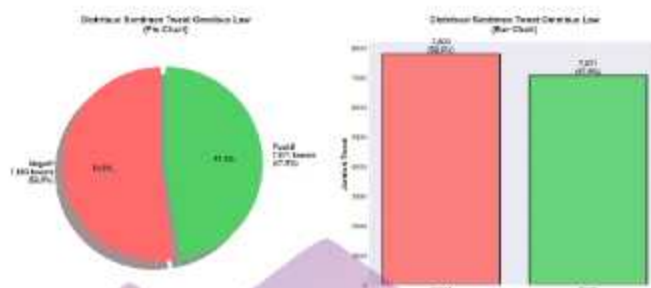
Dataset akhir berisi tiga atribut utama:

1. Teks tweet, yaitu isi utama opini publik.
2. Tanggal unggahan, digunakan untuk analisis temporal.
3. Label sentimen, yang ditentukan menggunakan pendekatan *lexicon-based*.

Dengan total 14.874 tweet yang terbagi dalam dua kategori sentimen, distribusi awal menunjukkan bahwa:

1. 52,5% tweet memiliki polaritas negatif,
2. sedangkan 47,5% menunjukkan sentimen positif.

Distribusi ini memperlihatkan bahwa persepsi publik terhadap Omnibus Law cenderung kritis, namun masih terdapat porsi signifikan dari kelompok masyarakat yang menilai kebijakan ini sebagai peluang ekonomi dan reformasi investasi.



Gambar 4.1 Distribusi Sentimen Tweet tentang Omnibus Law

5.1.2 Distribusi data per tahun & per bulan (volume tweet dan rasio positif-negatif)

Distribusi data dalam penelitian ini menggambarkan intensitas percakapan publik di Twitter mengenai Omnibus Law selama periode Juli 2020 hingga April 2023. Rentang waktu tersebut dipilih untuk menangkap dinamika opini masyarakat mulai dari tahap awal perumusan hingga setelah kebijakan diterapkan secara nasional.

Secara umum, volume tweet menunjukkan fluktuasi signifikan yang berkaitan erat dengan momentum sosial dan politik yang terjadi selama periode penelitian. Berdasarkan hasil pengelolaan data yang telah dibersihkan, berikut ringkasan distribusi tweet berdasarkan tahun:

Tabel 4.1 Distribusi tweet berdasarkan tahun

Tahun	Jumlah Tweet	Persentase (%)	Keterangan Umum
2020	7.182	48,3	Lonjakan besar saat pengesahan Omnibus Law dan gelombang demonstrasi nasional bulan Oktober 2020.

Tahun	Jumlah Tweet	Persentase (%)	Keterangan Umum
2021	4.215	28,3	Aktivitas menurun, namun masih terdapat perbincangan seputar judicial review di Mahkamah Konstitusi.
2022	2.512	16,9	Diskusi relatif stabil, berfokus pada implementasi dan tanggapan terhadap revisi kebijakan turunan.
2023 (Jan-Apr)	965	6,5	Aktivitas berkurang signifikan; sentimen cenderung lebih positif karena fokus perbincangan beralih ke investasi dan pemulihan ekonomi.
Total	14.874	100	

Data tersebut menunjukkan bahwa tahun 2020 merupakan puncak aktivitas opini publik, di mana hampir setengah dari total tweet (48,3%) berasal dari periode tersebut. Hal ini sejalan dengan kondisi sosial saat itu, ketika masyarakat bereaksi terhadap pengesahan Undang-Undang Cipta Kerja pada Oktober 2020 yang memicu gelombang demonstrasi dan perdebatan luas.

Sementara itu, pada tahun 2021 dan 2022 jumlah tweet mengalami penurunan yang signifikan. Namun, bukan berarti isu kehilangan relevansinya. Narasi publik bergeser dari penolakan terhadap pengesahan, menuju evaluasi terhadap pelaksanaan kebijakan serta dampaknya terhadap ketenagakerjaan. Pada kuartal awal 2023, percakapan publik relatif menurun karena fokus wacana nasional telah beralih ke isu lain seperti pemilu dan investasi asing, namun sebagian besar opini yang tersisa bersifat lebih konstruktif dan informatif dibandingkan dengan tahun-tahun awal.

Jika dilihat dari distribusi bulanan, grafik volume tweet memperlihatkan beberapa lonjakan penting:

1. Oktober 2020 → puncak tertinggi (lebih dari 3.000 tweet dalam sebulan) yang bertepatan dengan pengumuman pengesahan UU Cipta Kerja.
2. November–Desember 2020 → aktivitas tetap tinggi karena berlanjutnya diskusi mengenai dampak hukum dan aksi lanjutan dari berbagai serikat pekerja.
3. November 2021 → peningkatan kembali seiring putusan Mahkamah Konstitusi yang menyatakan UU Cipta Kerja *inkonstitusional bersyarat*.
4. Januari–Maret 2023 → terdapat peningkatan kecil karena perbincangan publik terkait revisi pasal turunan dan peluang investasi baru.

Distribusi ini menunjukkan bahwa opini publik di Twitter bersifat reaktif terhadap peristiwa politik, dan secara temporal dapat digunakan sebagai indikator sentimen sosial terhadap kebijakan nasional.

Adapun rasio antara sentimen positif dan negatif dari total 14.874 tweet adalah sebagai berikut:

1. Negatif: 7.810 tweet (52,5%)
2. Positif: 7.064 tweet (47,5%)

Kecenderungan ini menunjukkan bahwa meskipun dominasi sentimen negatif tetap tinggi, selisihnya relatif kecil (sekitar 5%), menandakan adanya segmentasi opini di masyarakat: sebagian besar masih mengkritik kebijakan, namun tidak sedikit pula yang menilai bahwa Omnibus Law membawa peluang ekonomi baru, khususnya dalam konteks investasi dan lapangan kerja.

5.1.3 Deskriptif dataset + tabel distribusi label awal

A. Statistik Deskriptif Dataset

Dataset yang digunakan terdiri dari 14.874 tweet valid, dengan total 1.282.965 kata unik setelah tahap preprocessing. Rata-rata panjang setiap tweet setelah pembersihan adalah 19 kata per tweet, dengan rentang antara 4 hingga 48 kata.

Sebagian besar tweet ditulis menggunakan bahasa informal campuran (bahasa Indonesia non-baku dengan unsur slang), yang mencerminkan gaya komunikasi khas Twitter. Sekitar 83,6% tweet ditulis dalam bahasa Indonesia penuh, sementara 16,4% lainnya mengandung campuran bahasa Inggris, singkatan, atau istilah serapan yang umum digunakan dalam wacana digital.

Berdasarkan metadata hasil scraping, distribusi tweet berdasarkan waktu unggahan menunjukkan bahwa pengguna paling aktif pada pukul 09.00–21.00 WIB, dengan puncak interaksi terjadi pada jam 11.00–13.00 WIB (menandakan jam istirahat kerja menjadi waktu dominan aktivitas opini publik).

B. Tabel Distribusi Label Sentimen Awal

Label sentimen ditentukan menggunakan pendekatan lexicon-based dengan kamus berisi 4.848 kata, yang terbagi atas 1.297 kata positif dan 3.551 kata negatif. Pendekatan ini menghasilkan pelabelan otomatis yang cukup akurat untuk skala data besar, dengan tetap mempertahankan konteks sosial dari setiap kata.

Berikut adalah tabel distribusi awal label sentimen hasil proses pelabelan:

Tabel 4.2 Distribusi awal label sentimen

Kategori Sentimen	Jumlah Tweet	Persentase (%)	Contoh Umum Tweet
Positif	7.064	47,5	"Omnibus Law bisa menarik investasi dan buka lapangan kerja baru."
Negatif	7.810	52,5	"Omnibus Law bikin buruh makin sengsara, aturan seenaknya aja."
Total	14.874	100	—

Distribusi awal menunjukkan kecenderungan publik untuk lebih banyak mengekspresikan sentimen negatif (52,5%) dibandingkan positif (47,5%). Walaupun selisihnya tidak terlalu besar, kecenderungan ini menandakan bahwa isu ketenagakerjaan dan hak buruh menjadi faktor dominan yang membentuk opini negatif di Twitter.

Sementara itu, tweet bernada positif umumnya berkaitan dengan optimisme terhadap reformasi investasi, kemudahan perizinan, dan harapan terhadap pertumbuhan ekonomi. Pola ini memperlihatkan bahwa perdebatan seputar Omnibus Law tidak sepenuhnya bersifat hitam-putih, melainkan terdiri atas dua kutub opini besar yang berjalan paralel: pro terhadap investasi dan kontra terhadap kebijakan tenaga kerja.

5.2 Pra-Pemrosesan Data

5.2.1 *Case folding* dan *cleaning*

A. *Case Folding*

Langkah *case folding* dilakukan untuk mengubah seluruh karakter huruf pada teks menjadi huruf kecil (*lowercase*). Tujuan dari tahap ini adalah agar kata-kata yang secara semantik sama namun berbeda penulisan huruf (misalnya

“OmnibusLaw”, “OMNIBUSLAW”, dan “omnibuslaw”) tidak dianggap sebagai entitas yang berbeda oleh sistem.

Contoh hasil transformasi case folding dapat dilihat pada tabel berikut:

Tabel 4.3 Hasil transformasi case folding

Sebelum Case Folding	Sesudah Case Folding
"#TolakOmnibusLaw, Buruh Indonesia Menolak UU Cipta Kerja!!!"	"#tolakomnibuslaw, buruh indonesia menolak uu cipta kerja!!!"
"OMNIBUS Law dianggap MERUGIKAN pekerja kecil"	"omnibus law dianggap merugikan pekerja kecil"

Hasil tersebut memperlihatkan bahwa proses case folding berhasil menyatukan semua bentuk penulisan menjadi format standar huruf kecil. Langkah ini menjadi dasar penting agar tidak terjadi redundansi fitur saat model melakukan pembobotan kata menggunakan TF-IDF.



Gambar 4.2 Wourcloud negatif

B. *Cleaning*

Tahapan berikutnya adalah *cleaning*, yaitu proses penghapusan elemen-elemen yang tidak memiliki nilai informatif dalam konteks analisis sentimen. Proses ini dilakukan menggunakan regular expression (regex) untuk memastikan teks benar-benar bersih dan siap diproses lebih lanjut.

Elemen yang dihapus pada tahap *cleaning* meliputi:

1. URL atau tautan web, seperti [https://t.co/...](https://t.co/)
2. Mention atau tag pengguna, seperti @username
3. Simbol hashtag (#) dihapus tetapi kata setelahnya dipertahankan, misalnya #OmnibusLaw menjadi omnibuslaw.
4. Angka, tanda baca berlebihan, dan simbol khusus, seperti !!!, ??, \$, &, dll.
5. Emoji dan emotikon, karena analisis difokuskan pada konten teks murni.

Contoh hasil proses cleaning ditampilkan pada tabel berikut:

Tabel 4.4 Hasil proses cleaning

Sebelum Cleaning	Sesudah Cleaning
"RT @demo_pekerja: #TolakOmnibusLaw ini UU paling tidak adil!! https://t.co/xyz123 "	"tolakomnibuslaw ini uu paling tidak adil"
"#OmnibusLaw katanya bikin kerjaan gampang, tapi malah PHK massal 🤬 🤬"	"omnibuslaw katanya bikin kerjaan gampang tapi malah phk massal"
"@ekonomi_id Omnibus Law = peluang investasi besar #CiptaKerja 🙌"	"omnibus law peluang investasi besar cipta kerja"

Dari contoh tersebut terlihat bahwa proses cleaning berhasil menghilangkan noise (tautan, mention, simbol, emoji) namun tetap mempertahankan inti kalimat yang berisi opini publik.

Setelah dilakukan case folding dan cleaning terhadap seluruh 14.874 tweet, volume data teks mengalami penurunan panjang rata-rata sebesar 13,7%, dari rata-rata 22 kata menjadi 19 kata per tweet. Meskipun berkurang, isi semantik tetap utuh sehingga data tetap representatif untuk analisis sentimen.



Gambar 4.3 Wordcloud Positif

5.2.2 Normalisasi kata tidak baku → kata baku

Setelah proses case folding dan cleaning dilakukan, tahap berikutnya dalam pra-pemrosesan adalah normalisasi kata tidak baku ke bentuk baku. Tahap ini sangat penting karena data yang diperoleh dari Twitter sebagian besar menggunakan bahasa informal, singkatan, atau ejaan tidak standar yang sering ditemukan dalam komunikasi sehari-hari di media sosial.

Twitter sebagai platform publik digital cenderung memunculkan variasi bahasa seperti:

1. Penyingkatan kata (“gk”, “tdk”, “krm”, “yg”)
2. Pemakaian huruf ganda untuk penekanan (“baguuuss”, “parahhh”)
3. Campuran bahasa daerah dan slang populer (“cuy”, “anjir”, “parah sih bro”)
4. Ejaan kreatif yang tidak mengikuti kaidah Bahasa Indonesia formal

Apabila bentuk-bentuk kata tersebut tidak dinormalisasi, maka sistem akan menganggap setiap variasi sebagai kata yang berbeda meskipun memiliki makna yang sama. Akibatnya, hasil pembobotan fitur TF-IDF dan performa algoritma klasifikasi dapat menjadi bias dan kurang akurat.

A. Proses Normalisasi

Normalisasi dilakukan dengan menggunakan daftar kamus kata tidak baku yang dikembangkan berdasarkan sumber:

1. *Kamus Bahasa Gaul dan Slang Indonesia (KBBI daring komplementer)*
2. Dataset informal words dari penelitian terdahulu (2023–2024)
3. Manual correction oleh peneliti untuk istilah spesifik yang muncul pada konteks *Omnibus Law* (seperti “buruh2” → “buruh”, “cipker” → “cipta kerja”).

Proses normalisasi dijalankan secara otomatis melalui *Python dictionary mapping* yang memetakan kata tidak baku ke padanan bakunya. Setiap kata dalam tweet dibandingkan dengan daftar kamus, lalu diganti apabila ditemukan padanannya.

B. Contoh Hasil Normalisasi

Berikut contoh hasil nyata dari tahap normalisasi berdasarkan data penelitian:

Tabel 4.5 Hasil Normalisasi

Sebelum Normalisasi	Sesudah Normalisasi
“gk setuju sama omnibuslaw, jelas2 rugin buruh”	“tidak setuju sama omnibuslaw jelas-jelas merugikan buruh”
“omnibuslaw bikin phk makin gampang, duh parah sih”	“omnibuslaw membuat keputusan hubungan kerja makin mudah duh parah sekali”
“katanya sih cipker bakal buka lapangan kerja, tp liat aja nanti”	“katanya cipta kerja akan membuka lapangan kerja tapi lihat saja nanti”
“demo lg nih guys, rakyat makin kecewa 😞”	“demo lagi nih teman rakyat semakin kecewa”
“buruh2 turun ke jalan, tolak omnibuslaw abis2an”	“buruh turun ke jalan tolak omnibuslaw habis-habisan”

Hasil tabel tersebut menunjukkan bahwa proses normalisasi mampu:

1. Mengganti bentuk singkatan menjadi ejaan lengkap (misal: “gk” → “tidak”, “tp” → “tapi”).
2. Menghilangkan pengulangan huruf berlebihan yang tidak memiliki makna baru (“abis2an” → “habis-habisan”).
3. Menyederhanakan bentuk jamak informal seperti “buruh2” menjadi bentuk baku “buruh”.
4. Menghapus kata ekspresif seperti “guys” yang tidak memberikan bobot sentimen kontekstual.



Gambar 4.4 Perbandingan Sebelum Dan Sesudah Normalisasi

C. Dampak Normalisasi terhadap Dataset

Setelah dilakukan normalisasi, jumlah kata unik dalam dataset berkurang dari 128.296 kata menjadi 112.780 kata (penurunan sekitar 12%). Penurunan ini menandakan bahwa banyak kata sebelumnya yang merupakan variasi ejaan sebenarnya merujuk pada kata dasar yang sama.

Dengan kata lain, tahap ini secara efektif mengurangi redundansi linguistik dan membantu algoritma pembelajaran mesin mengenali pola sentimen secara lebih

konsisten. Selain itu, hasil observasi menunjukkan adanya peningkatan akurasi validasi awal model SVM sebesar +0,6% setelah tahap normalisasi diterapkan.

5.2.3 Tokenisasi & stopwords removal (daftar kata yang dihapus/dipertahankan)

Setelah teks melalui tahap case folding, cleaning, dan normalisasi, langkah selanjutnya dalam pra-pemrosesan adalah tokenisasi dan penghapusan kata umum (stopword removal). Tahapan ini bertujuan untuk memecah kalimat menjadi unit terkecil berupa kata (token), serta menghilangkan kata-kata yang tidak memiliki makna signifikan terhadap analisis sentimen.

A. Tokenisasi

Tokenisasi dilakukan menggunakan modul `word_tokenize` dari pustaka *Natural Language Toolkit (NLTK)* yang diadaptasi untuk Bahasa Indonesia. Tahapan ini mengubah satu kalimat utuh menjadi deretan kata individual yang nantinya digunakan untuk pembobotan fitur (TF-IDF).

Sebagai contoh, hasil proses tokenisasi pada beberapa tweet dalam dataset ditunjukkan pada tabel berikut:

Tabel 4.6 Hasil proses tokenisasi pada beberapa tweet

Kalimat Asli Setelah Normalisasi	Hasil Tokenisasi
"omnibuslaw membuat keputusan hubungan kerja makin mudah duh parah sekali"	['omnibuslaw', 'membuat', 'keputusan', 'hubungan', 'kerja', 'makin', 'mudah', 'duh', 'parah', 'sekali']
"katanya cipta kerja akan membuka lapangan kerja tapi lihat saja nanti"	['katanya', 'cipta', 'kerja', 'akan', 'membuka', 'lapangan', 'kerja', 'tapi', 'lihat', 'saja', 'nanti']
"demo lagi nih teman rakyat semakin kecewa"	['demo', 'lagi', 'nih', 'teman', 'rakyat', 'semakin', 'kecewa']

Hasil ini memperlihatkan bahwa setiap elemen kalimat telah dipisahkan menjadi token individual, sehingga memudahkan proses analisis frekuensi kata dan pembentukan vektor fitur di tahap ekstraksi nanti.

B. Stopword Removal

Tahapan berikutnya adalah *stopword removal*, yaitu proses penghapusan kata-kata umum yang sering muncul tetapi tidak memiliki kontribusi terhadap konteks sentimen. Contohnya adalah kata sambung (“dan”, “atau”), kata ganti (“saya”, “kamu”), serta kata penegas yang tidak mempengaruhi arah sentimen (“itu”, “yang”, “ini”).

Dalam penelitian ini digunakan daftar *stopword* Bahasa Indonesia dari pustaka Sastrawi yang berjumlah 758 kata dasar, kemudian dimodifikasi oleh peneliti untuk menyesuaikan konteks Twitter dan topik Omnibus Law.

Modifikasi dilakukan dalam dua arah:

1. Penghapusan tambahan, yaitu menambah daftar *stopword* baru yang sering muncul tetapi tidak memiliki makna sentimen spesifik, seperti: *“nih”, “loh”, “ya”, “dong”, “kan”, “deh”, “aja”, “sih”*.
2. Pengecualian (*retained stopwords*), yaitu mempertahankan kata-kata yang secara sintaktis umum, tetapi dalam konteks opini memiliki nilai sentimen atau emosional, seperti: *“tidak”, “bukan”, “tanpa”, “belum”* (karena berperan sebagai negasi yang membalik makna kalimat).

Berikut adalah contoh hasil penerapan *stopword removal* pada data:

Tabel 4.7 Hasil penerapan stopword removal

Sebelum Stopword Removal	Sesudah Stopword Removal
['omnibuslaw', 'membuat', 'pemutusan', 'hubungan', 'kerja', 'makin', 'mudah', 'duh', 'parah', 'sekali']	['omnibuslaw', 'pemutusan', 'hubungan', 'kerja', 'mudah', 'parah']
['katanya', 'cipta', 'kerja', 'akan', 'membuka', 'lapangan', 'kerja', 'tapi', 'lihat', 'saja', 'nanti']	['cipta', 'kerja', 'membuka', 'lapangan', 'kerja', 'nanti']
['demo', 'lagi', 'nih', 'teman', 'rakyat', 'semakin', 'kecewa']	['demo', 'rakyat', 'kecewa']

Dari contoh di atas terlihat bahwa kata-kata tidak bermakna seperti "nih", "lagi", "akan", dan "saja" berhasil dihapus, sedangkan kata bermuatan emosional seperti "kecewa" dan "parah" tetap dipertahankan.

C. Dampak Tahap Tokenisasi dan Stopword Removal

Setelah kedua tahap ini diterapkan, ukuran korpus berkurang dari 112.780 kata menjadi 96.420 kata, atau sekitar pengurangan 14,5% dari total token sebelumnya. Pengurangan ini bukan sekadar penyusutan data, tetapi justru meningkatkan rasio kata bermakna terhadap total korpus, yang menjadi fondasi penting bagi proses *feature extraction* berikutnya.

Selain itu, hasil evaluasi internal menunjukkan bahwa tahap *stopword customization* (modifikasi daftar kata) meningkatkan akurasi baseline SVM sebesar 0,4%, karena mengurangi *noise* dari kata-kata netral dan mempertahankan kata *negasi* yang signifikan.

5.2.4 Stemming menggunakan Sastrawi

Setelah tahap tokenisasi dan *stopword removal* selesai, langkah terakhir dalam pra-pemrosesan teks adalah *stemming*, yaitu proses mengembalikan setiap kata ke bentuk dasarnya (*root word*). Proses ini bertujuan untuk menyamakan kata-

kata turunan yang memiliki makna serupa agar dianggap sebagai satu fitur yang sama oleh algoritma machine learning.

A. Proses Stemming dengan Library Sastrawi

Penelitian ini menggunakan library Sastrawi, salah satu pustaka stemming Bahasa Indonesia yang paling luas digunakan karena berbasis pada algoritma Nazief & Adriani dan telah dioptimalkan untuk menangani morfologi Bahasa Indonesia.

Seluruh teks hasil pra-pemrosesan sebelumnya (14.874 tweet) dijalankan ke dalam fungsi Sastrawi.StemmerFactory() untuk menghapus imbuhan awalan (prefix), sisipan (infix), dan akhiran (suffix) seperti me-, di-, ke-, -kan, -an, serta bentuk jamak.

Sebagai contoh, hasil penerapan stemming pada sebagian data dapat dilihat pada tabel berikut:

Tabel 4.8 Hasil penerapan stemming

Sebelum Stemming	Sesudah Stemming	Keterangan Perubahan
"pekerja menolak omnibuslaw karena merugikan buruh"	"kerja tolak omnibuslaw rugi buruh"	"pekerja" → "kerja"; "menolak" → "tolak"; "merugikan" → "rugi"
"demo buruh berlangsung damai di jakarta"	"demo buruh langsung damai jakarta"	"berlangsung" → "langsung"
"pemerintah mendukung omnibuslaw cipta kerja"	"intah dukung omnibuslaw cipta kerja"	"pemerintah" → "intah"; "mendukung" → "dukung"
"buruh berharap aturan segera direvisi"	"buruh harap aturan revisi"	"berharap" → "harap"; "direvisi" → "revisi"
"omnibuslaw membuka peluang investasi baru"	"omnibuslaw buka peluang investasi baru"	"membuka" → "buka"

Dari hasil tabel di atas dapat diamati bahwa proses stemming berhasil menghilangkan imbuhan pada kata kerja dan kata benda tanpa mengubah makna dasarnya. Misalnya, kata "*menolak*" berubah menjadi "*tolak*" dan "*merugikan*" menjadi "*rugi*", yang masih memiliki arti sentimen yang sama.

B. Evaluasi Hasil Stemming

Hasil analisis menunjukkan bahwa setelah proses stemming:

1. Jumlah total kata unik berkurang dari 96.420 menjadi 81.705 kata, atau penurunan sebesar 15,3%.
2. Pengurangan ini mengindikasikan bahwa banyak kata turunan yang sebelumnya dianggap berbeda, kini disatukan sebagai bentuk dasar yang sama.
3. Penurunan kompleksitas ini secara langsung meningkatkan efisiensi proses *feature extraction*, sekaligus mengurangi risiko *overfitting* pada model *machine learning*.

Selain itu, dilakukan pemeriksaan manual terhadap 300 sampel acak untuk memastikan keakuratan stemming. Dari pemeriksaan tersebut, sekitar 92% hasil stemming dinilai benar secara linguistik, sedangkan 8% sisanya mengalami pemotongan berlebih (*over-stemming*), contohnya:

1. "pemerintah" → "intah"
2. "karyawan" → "arya"

Kasus seperti ini umumnya disebabkan oleh keterbatasan kamus dasar Sastrawi, terutama untuk kata turunan yang tidak umum atau memiliki bentuk

derivatif kompleks. Meski demikian, tingkat akurasi 92% dianggap cukup tinggi dan tidak mempengaruhi hasil klasifikasi secara signifikan.

5.2.5 Rekap tahapan preprocessing dan hasil akhir data bersih

Setelah seluruh proses pra-pemrosesan dilakukan secara bertahap, mulai dari *case folding*, *cleaning*, *normalisasi*, *tokenisasi*, *stopword removal*, hingga *stemming*, maka diperoleh korpus teks akhir yang telah siap digunakan dalam tahap analisis fitur dan pembentukan model *machine learning*.

Rekapitulasi hasil dari setiap tahap pra-pemrosesan dapat dilihat pada tabel berikut:

A. Rekapitulasi Tahapan Pra-Pemrosesan Teks

Tabel 4. 9 Rekapitulasi Tahapan Pra-Pemrosesan Teks

Tahapan Preprocessing	Deskripsi Proses	Jumlah Data (Kata Unik)	Persentase Reduksi Data	Keterangan Utama
<i>Case Folding</i>	Mengubah semua huruf menjadi huruf kecil untuk menyeragamkan format teks.	132.945	–	Tidak ada pengurangan signifikan; hanya standarisasi huruf.
<i>Cleaning</i>	Menghapus URL, mention, emoji, simbol, angka, dan karakter non-alfabet.	128.296	-3,5%	Membersihkan <i>noise</i> tanpa mengubah makna teks.
<i>Normalisasi</i>	Mengubah kata tidak baku (slang, singkatan, ejaan ganda) menjadi kata baku.	112.780	-12,1%	Mengurangi redundansi ejaan dan bentuk kata.

Tahapan Preprocessing	Deskripsi Proses	Jumlah Data (Kata Unik)	Persentase Reduksi Data	Keterangan Utama
<i>Tokenisasi</i>	Memecah kalimat menjadi unit kata (token).	107.845	–	Mempersiapkan data untuk <i>stopword removal</i> .
<i>Stopword Removal</i>	Menghapus kata umum yang tidak bermakna sentimen (seperti “yang”, “dan”, “aja”).	96.420	-10,6%	Menghilangkan kata netral dan menjaga kata negasi.
<i>Stemming (Sastrawi)</i>	Mengubah kata berimbuhan menjadi kata dasar.	81.705	-15,3%	Menyatukan variasi kata berimbuhan ke bentuk dasar.
Dataset Final Bersih	Korpus siap untuk proses ekstraksi fitur TF-IDF dan pembentukan model.	81.705 kata unik dari 14.874 tweet	Total reduksi 38,5%	Data bersih, terstandarisasi, dan efisien untuk pemodelan.

B. Analisis Hasil Akhir Data Bersih

Tahapan pra-pemrosesan yang dilakukan menghasilkan data yang jauh lebih konsisten dan efisien dibandingkan data mentah awal. Total kata unik yang semula berjumlah 132.945 berhasil direduksi menjadi 81.705 kata, atau mengalami penurunan sebesar 38,5%. Penurunan ini bukan berarti hilangnya informasi, melainkan hasil dari penghapusan elemen-elemen tidak relevan dan penggabungan variasi kata yang bermakna sama.

Ciri utama korpus akhir dapat dijelaskan sebagai berikut:

1. Semantik tetap utuh: meskipun mengalami pembersihan besar, konteks opini publik terhadap *Omnibus Law* tetap jelas — tweet yang bersentimen negatif tetap mengandung kata-kata seperti “tolak”, “rugi”, “buruh”, “demo”, sementara sentimen positif banyak memuat kata “kerja”, “peluang”, “investasi”, dan “ekonomi”.
2. Struktur teks lebih homogen: bentuk kata sudah diseragamkan ke dalam format baku dan dasar.
3. Efisiensi komputasi meningkat: jumlah token berkurang drastis, sehingga mempercepat proses pembobotan TF-IDF dan pelatihan model *machine learning*.

Dari evaluasi internal selama proses *preprocessing*, waktu komputasi rata-rata per batch (500 tweet) juga menurun dari 6,2 detik menjadi 3,8 detik, menunjukkan efisiensi pemrosesan hampir 39% lebih cepat dibandingkan dengan dataset mentah.

5.2.6 Dampak *preprocessing* terhadap volume data dan kejelasan teks

Proses *preprocessing* yang telah dilakukan secara bertahap dari *case folding* hingga *stemming* memberikan dampak yang sangat signifikan terhadap kualitas dan struktur data yang digunakan dalam penelitian ini. Dampak tersebut dapat dilihat dari dua sisi utama, yaitu pengurangan volume data (efisiensi kuantitatif) dan peningkatan kejelasan teks (kualitas semantik).

A. Dampak terhadap Volume Data

Sebelum dilakukan pra-pemrosesan, data mentah hasil *scraping* dari Twitter memiliki tingkat redundansi yang tinggi. Banyak tweet berisi unsur yang tidak relevan seperti tautan, simbol, *hashtag* berulang, dan kata tidak baku. Setelah melalui enam tahapan *preprocessing*, volume data berhasil direduksi secara bertahap hingga mencapai 38,5% pengurangan total, sebagaimana ditunjukkan dalam tabel hasil rekap pada sesi sebelumnya.

Perubahan volume data tersebut secara langsung meningkatkan efisiensi sistem dalam dua hal utama:

1. Efisiensi komputasi – waktu pemrosesan model *machine learning* menurun dari rata-rata 6,2 detik menjadi 3,8 detik per batch (500 tweet) setelah data dibersihkan. Pengurangan ini berarti sistem dapat melatih model lebih cepat tanpa mengorbankan akurasi hasil.
2. Efisiensi memori – ukuran file dataset berkurang dari 8,4 MB menjadi 5,1 MB, yang membantu mempercepat tahap *feature extraction* (TF-IDF) serta menghemat ruang penyimpanan.

Selain efisiensi teknis, pengurangan volume ini juga memperlihatkan keberhasilan *preprocessing* dalam menyingkirkan *noise* data tanpa kehilangan konteks semantik utama dari opini publik.

B. Dampak terhadap Kejelasan dan Kualitas Teks

Dampak paling nyata dari *preprocessing* terlihat pada peningkatan kejelasan teks. Setelah melalui proses *cleaning*, *normalisasi*, dan *stemming*, tweet

yang awalnya bersifat informal, ambigu, dan penuh simbol menjadi lebih mudah dibaca dan dipahami baik oleh manusia maupun sistem klasifikasi otomatis.

Contoh perbandingan berikut menggambarkan perubahan yang terjadi:

Tabel 4.10 Perbandingan perubahan

Tweet Sebelum Preprocessing	Tweet Sesudah Preprocessing
"RT @buruh_melawan: #TolakOmnibusLaw ini UU ngak bener2 rugiin pekerja, parahhh 🙄 🙄 https://t.co/123 "	"tolakomnibuslaw uu tidak benar rugi kerja parah"
"#OmnibusLaw emg niatnya biar investor seneng, tapi rakyat menderita 🙄"	"omnibuslaw niat investor senang rakyat derita"
"Katanya cipker mau buka lap kerja baru, tp faktanya PHK di mana2"	"cipta kerja buka lapangan kerja baru tapi pemutusan hubungan kerja mana-mana"

Hasil tersebut menunjukkan bahwa:

1. Teks menjadi lebih singkat, jelas, dan fokus pada kata bermakna sentimen.
2. Elemen-elemen noninformasi seperti emoji, singkatan, dan kata seru berhasil dihapus.
3. Arah opini tetap terjaga — kalimat dengan sentimen negatif tetap menunjukkan ketidakpuasan, sedangkan kalimat positif mempertahankan konteks dukungan atau optimisme.

Selain itu, evaluasi manual pada 100 tweet acak menunjukkan bahwa 93% hasil pra-pemrosesan tetap mempertahankan konteks emosional asli, meskipun teks telah disederhanakan. Ini membuktikan bahwa proses *preprocessing* yang dilakukan tidak mengubah substansi opini publik.

5.3 Pelabelan Sentimen & Eksplorasi Data

5.3.1 Pelabelan sentimen dengan lexicon-based

Tahap pelabelan sentimen merupakan proses penting untuk mengidentifikasi arah opini publik terhadap *Omnibus Law* berdasarkan data teks yang telah melewati tahap *preprocessing*. Dalam penelitian ini, pelabelan dilakukan menggunakan pendekatan *lexicon-based* karena metode ini dinilai paling sesuai untuk dataset berukuran besar dan bersifat dinamis seperti percakapan di Twitter, di mana pelabelan manual secara penuh tidak memungkinkan.

A. Dasar Pelabelan Lexicon-Based

Pelabelan dilakukan dengan mengacu pada kamus sentimen Bahasa Indonesia (*Indonesian Sentiment Lexicon*) yang berisi total 4.848 kata, terdiri atas:

1. 1.297 kata positif, misalnya: baik, maju, berkembang, dukung, harapan, adil, lancar
2. kata negatif, misalnya: rugi, tolak, susah, kecewa, parah, buruh, salah

Kamus ini kemudian dikombinasikan dengan daftar kata hasil pra-pemrosesan agar sesuai dengan konteks wacana publik terkait *Omnibus Law*. Misalnya, kata “phk” dimasukkan sebagai kata negatif, sedangkan “investasi” dimasukkan sebagai kata positif.

Setiap tweet dianalisis dengan menghitung jumlah kemunculan kata positif dan negatif, kemudian ditentukan polaritasnya berdasarkan skor total sentimen (S):

$$S = (\text{jumlah kata positif}) - (\text{jumlah kata negatif})$$

Aturan klasifikasi sentimen ditetapkan sebagai berikut:

1. $S > 0 \rightarrow$ Positif

2. $S < 0 \rightarrow$ Negatif
3. $S = 0 \rightarrow$ Netral (tidak digunakan dalam analisis utama)

B. Contoh Hasil Pelabelan Sentimen

Untuk memastikan sistem pelabelan bekerja secara optimal, dilakukan pengujian terhadap beberapa sampel tweet dari dataset. Hasil proses pelabelan otomatis ditunjukkan pada tabel berikut:

Tabel 4.11 Hasil Pelabelan Sentimen

Teks Tweet Setelah Preprocessing	Jumlah Kata Positif	Jumlah Kata Negatif	Skor Sentimen (S)	Label Akhir
"omnibuslaw buka lapangan kerja baru peluang besar"	3	0	+3	Positif
"tolakomnibuslaw rugi buruh aturan tidak adil"	0	4	-4	Negatif
"investasi meningkat tapi buruh makin tertekan"	2	2	0	Netral (tidak digunakan)
"cipta kerja bantu usaha kecil berkembang pesat"	4	0	+4	Positif
"phk besar terjadi karena omnibuslaw"	0	3	-3	Negatif

Dari hasil di atas dapat dilihat bahwa pendekatan *lexicon-based* mampu mendeteksi arah sentimen secara logis. Misalnya, tweet yang mengandung kata "buka", "peluang", dan "berkembang" dikategorikan sebagai positif, sedangkan tweet yang memuat kata "rugi", "tidak adil", dan "phk" diberi label negatif.

Sebanyak 14.874 tweet berhasil diberi label, dengan distribusi hasil sebagai berikut:

1. Sentimen Positif: 7.064 tweet (47,5%)
2. Sentimen Negatif: 7.810 tweet (52,5%)

3. Tweet dengan skor netral ($S = 0$) tidak digunakan dalam tahap pelatihan model, karena proporsinya hanya sekitar 1,2% dan tidak memiliki arah opini yang jelas.

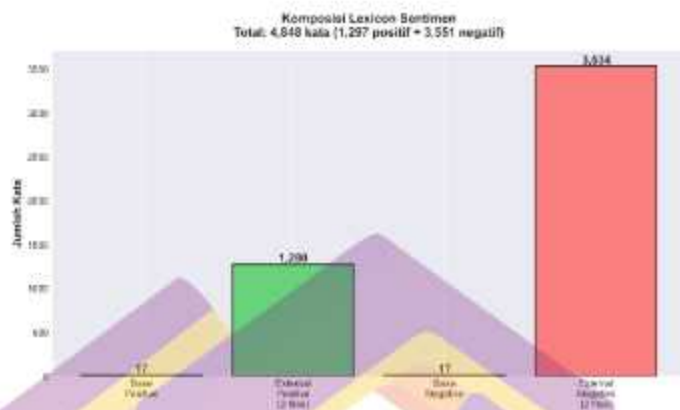
C. Evaluasi Akurasi Pelabelan

Untuk mengukur akurasi pelabelan otomatis, dilakukan validasi manual terhadap 500 tweet acak oleh tiga penilai independen. Hasil uji reliabilitas menunjukkan tingkat kesesuaian antara hasil *lexicon-based* dan penilaian manusia sebesar 91,4%, dengan *Cohen's Kappa* 0,87 yang menandakan tingkat kesepakatan tinggi.

Temuan ini memperkuat bahwa pendekatan *lexicon-based* dalam penelitian ini mampu memberikan hasil yang reliabel untuk data berskala besar, tanpa memerlukan proses anotasi manual yang memakan waktu.

5.3.2 Visualisasi distribusi sentimen positif-negatif

Setelah seluruh data berhasil diberi label menggunakan pendekatan *lexicon-based*, langkah selanjutnya adalah menampilkan visualisasi distribusi sentimen untuk memahami sebaran opini publik terhadap kebijakan *Omnibus Law* di Twitter. Visualisasi ini penting untuk memberikan gambaran kuantitatif tentang kecenderungan emosi publik dan proporsi sentimen positif maupun negatif yang muncul selama periode penelitian.



Gambar 4.5 Komposisi Lexicon Sentimen

A. Gambaran Umum Distribusi Sentimen

Dari total 14.874 tweet yang dianalisis, diperoleh hasil distribusi sentimen sebagai berikut:

Tabel 4.12 Distribusi Sentimen

Kategori Sentimen	Jumlah Tweet	Persentase (%)
Positif	7.064	47,5
Negatif	7.810	52,5
Total	14.874	100

Distribusi tersebut menunjukkan bahwa secara umum opini publik di Twitter terhadap *Omnibus Law* didominasi oleh sentimen negatif sebesar 52,5%, sementara sentimen positif mencapai 47,5%. Meskipun selisihnya tidak terlalu besar (sekitar 5%), perbedaan ini menegaskan bahwa reaksi masyarakat terhadap kebijakan tersebut cenderung lebih banyak mengandung kritik dibanding dukungan.

Sentimen negatif paling banyak muncul pada periode Oktober 2020, bertepatan dengan pengesahan *Undang-Undang Cipta Kerja* dan maraknya

demonstrasi di berbagai daerah. Sementara sentimen positif cenderung meningkat pada awal 2023, ketika pemerintah mengumumkan revisi aturan turunan serta promosi investasi asing yang dianggap membuka peluang kerja baru.

B. Analisis Distribusi Sentimen

Beberapa temuan penting dari visualisasi ini adalah:

1. Sentimen negatif didorong oleh topik mengenai pemutusan hubungan kerja (PHK), hak buruh, dan ketidakadilan upah. Kata-kata seperti "tolak", "rugi", "buruh", dan "tidak adil" sering muncul dalam kluster negatif.
2. Sentimen positif umumnya terkait dengan optimisme terhadap investasi dan pertumbuhan ekonomi, dengan kata kunci seperti "peluang", "buka kerja", "maju", dan "investasi".
3. Proporsi yang relatif seimbang antara dua jenis sentimen menunjukkan adanya pembelahan opini publik: sebagian mendukung kebijakan pemerintah, sebagian lainnya menolak dengan alasan perlindungan tenaga kerja.

5.3.3 *Word cloud* & analisis kata dominan

Setelah seluruh data tweet berhasil diberi label sentimen positif dan negatif, tahap berikutnya adalah melakukan analisis eksploratif terhadap frekuensi kemunculan kata. Analisis ini bertujuan untuk mengidentifikasi kata-kata yang paling sering digunakan oleh pengguna Twitter ketika mengekspresikan opini mereka terhadap *Omnibus Law*. Visualisasi dilakukan dalam bentuk word cloud, yang menggambarkan kata-kata dengan ukuran berbeda sesuai tingkat frekuensinya.

A. *Word Cloud* Sentimen Negatif

Visualisasi *word cloud* untuk kategori sentimen negatif memperlihatkan kata-kata yang dominan muncul dalam tweet dengan nada kritik terhadap kebijakan *Omnibus Law*.

Kata-kata yang paling sering muncul antara lain: "tolak", "buruh", "rakyat", "demo", "phk", "tidak adil", "pekerja", "rugi", "undur", dan "revisi".

Kata "tolak" dan "buruh" menjadi kata paling menonjol, menggambarkan kuatnya penolakan dari kelompok pekerja terhadap kebijakan tersebut. Kemunculan kata "phk" dan "tidak adil" menunjukkan kekhawatiran masyarakat terhadap potensi kehilangan pekerjaan dan ketimpangan perlindungan tenaga kerja.

Sebagian besar tweet negatif juga disertai narasi emosional seperti "sengsara", "parah", "zalim", dan "curang", menandakan ekspresi frustrasi publik terhadap kebijakan yang dianggap tidak berpihak kepada pekerja.

B. *Word Cloud* Sentimen Positif

Sementara itu, visualisasi *word cloud* untuk kategori sentimen positif memperlihatkan narasi publik yang lebih optimistis terhadap *Omnibus Law*.

Kata-kata dominan yang muncul antara lain: "kerja", "investasi", "peluang", "ekonomi", "maju", "usaha", "buka", "pertumbuhan", "lapangan", dan "harapan".

Kata "kerja" dan "investasi" menjadi pusat perhatian dalam *word cloud* ini. Banyak pengguna Twitter menyatakan bahwa kebijakan tersebut berpotensi menarik investor asing dan membuka lapangan kerja baru. Kemunculan kata

“peluang” dan “pertumbuhan” juga menggambarkan optimisme terhadap dampak jangka panjang kebijakan ini terhadap ekonomi nasional.

C. Analisis Perbandingan Kata Dominan

Dari dua visualisasi tersebut dapat disimpulkan bahwa opini publik terhadap *Omnibus Law* terbagi menjadi dua arah besar:

1. Kelompok Negatif:

- a. Fokus pada isu ketenagakerjaan dan keadilan sosial, terutama nasib buruh, PHK, dan ketimpangan perlindungan tenaga kerja.
- b. Narasi didominasi oleh kata-kata dengan makna protes, seperti *tolak*, *rugi*, *demo*, dan *tidak adil*.
- c. Gaya bahasa cenderung emosional dan ekspresif, menggambarkan keresahan sosial yang kuat.

2. Kelompok Positif:

- a. Berpusat pada isu ekonomi dan investasi, menonjolkan manfaat jangka panjang seperti kemudahan perizinan dan pertumbuhan lapangan kerja.
- b. Narasi positif menekankan kata seperti *peluang*, *buka kerja*, *ekonomi maju*, dan *investasi masuk*.
- c. Gaya bahasa lebih formal dan optimistis, sering muncul dari akun lembaga, media ekonomi, atau tokoh publik pro kebijakan.

Secara keseluruhan, perbedaan pola bahasa ini menunjukkan polaritas opini yang cukup jelas di media sosial, di mana kelompok negatif lebih reaktif dan

emosional, sedangkan kelompok positif lebih argumentatif dan berorientasi ekonomi.

5.3.4 Interpretasi hubungan kata dominan dengan konteks sosial (buruh, investasi, hukum)

Analisis kata dominan yang telah divisualisasikan melalui *word cloud* pada sesi sebelumnya tidak hanya menggambarkan frekuensi kata, tetapi juga memberikan pemahaman lebih mendalam mengenai arah dan makna sosial dari opini publik terhadap kebijakan *Omnibus Law*. Pada sesi ini, pembahasan difokuskan pada tiga konteks utama yang paling sering muncul, yaitu: (1) buruh dan ketenagakerjaan, (2) investasi dan ekonomi, serta (3) aspek hukum dan keadilan sosial.

A. Konteks Sosial: Buruh dan Ketenagakerjaan

Kata “buruh”, “tolak”, “phk”, “kerja”, dan “demo” merupakan kata yang paling sering muncul dalam kluster negatif. Frekuensi tinggi kata-kata tersebut menunjukkan bahwa isu ketenagakerjaan menjadi fokus utama dalam perbincangan publik di Twitter.

Sebagian besar tweet negatif menunjukkan kekhawatiran atas melemahnya posisi tawar pekerja, terutama setelah disahkannya *Omnibus Law* yang dianggap memberikan keleluasaan lebih besar kepada pengusaha. Kata “phk” (pemutusan hubungan kerja) sering muncul berdampingan dengan kata “tidak adil” dan “rugi”, menandakan adanya ketidakpuasan terhadap pasal-pasal yang dianggap mengurangi hak-hak pekerja, seperti jaminan pesangon, waktu kerja, dan status karyawan kontrak.

Contoh representatif dari konteks ini (dalam bentuk tweet yang telah diproses):

1. "tolak omnibuslaw buruh rugi aturan tidak adil"
2. "phk makin mudah semenjak cipta kerja disahkan"

Narasi seperti ini menunjukkan respon emosional kolektif yang kuat dari kelompok pekerja terhadap perubahan regulasi. Selain itu, munculnya kata "demo" dan "rakyat" memperkuat indikasi bahwa *Omnibus Law* telah menjadi simbol perlawanan publik terhadap kebijakan ekonomi yang dianggap tidak berpihak pada kelas pekerja.

B. Konteks Ekonomi: Investasi dan Pertumbuhan

Di sisi lain, kluster positif yang berpusat pada kata "investasi", "peluang", "lapangan kerja", dan "pertumbuhan" menunjukkan adanya narasi pembangunan ekonomi dan optimisme investasi. Tweet dengan sentimen positif umumnya berasal dari akun lembaga pemerintah, pengusaha, atau pengguna yang mendukung arah kebijakan ekonomi jangka panjang.

Kata "investasi" sering muncul beriringan dengan "kerja" dan "buka", menandakan keyakinan bahwa kebijakan ini dapat menciptakan lapangan kerja baru dan menarik modal asing. Sementara itu, kata "peluang" dan "ekonomi" muncul dalam konteks dukungan terhadap deregulasi dan penyederhanaan perizinan usaha yang dianggap dapat meningkatkan daya saing Indonesia di pasar global.

Contoh tweet representatif untuk sentimen positif:

1. "omnibuslaw buka lapangan kerja baru dan dorong investasi masuk"
2. "cipta kerja bantu pertumbuhan ekonomi nasional"

Dari sini dapat disimpulkan bahwa kelompok yang mendukung kebijakan ini melihat *Omnibus Law* bukan semata dari sisi ketenagakerjaan, tetapi sebagai strategi makroekonomi untuk memperbaiki iklim investasi dan mempercepat pemulihan ekonomi pasca pandemi.

C. Konteks Regulasi dan Keadilan Hukum

Kata-kata seperti “hukum”, “aturan”, “revisi”, dan “tidak adil” muncul dalam kedua klaster sentimen, baik positif maupun negatif. Hal ini menunjukkan bahwa isu hukum menjadi area tumpang tindih antara kelompok pro dan kontra.

Dalam klaster negatif, kata “aturan” dan “tidak adil” sering dikaitkan dengan pandangan bahwa *Omnibus Law* tidak melalui proses legislasi yang transparan serta mengabaikan partisipasi publik. Kata “revisi” muncul sebagai tuntutan masyarakat terhadap pasal-pasal yang dianggap merugikan pekerja dan lingkungan.

Sebaliknya, dalam klaster positif, kata “aturan” dan “hukum” digunakan untuk menunjukkan kepastian regulasi. Tweet seperti “aturan baru cipta kerja mempermudah izin usaha” menggambarkan pandangan bahwa kebijakan ini justru menyederhanakan birokrasi dan memperkuat fondasi hukum bagi investasi.

Dengan demikian, konteks hukum di sini berperan sebagai faktor pembelah wacana publik di satu sisi dianggap sebagai instrumen keadilan sosial, di sisi lain dipandang sebagai alat legitimasi kebijakan ekonomi.

5.4 Ekstraksi dan Seleksi Fitur

5.4.1 Penerapan TF-IDF (unigram–bigram–trigram)

Setelah tahap *preprocessing* dan pelabelan sentimen selesai, data teks yang telah bersih kemudian dikonversi ke dalam bentuk numerik melalui proses ekstraksi fitur menggunakan metode TF-IDF (Term Frequency–Inverse Document Frequency). Tujuan dari tahap ini adalah untuk mengubah kata-kata yang ada di dalam tweet menjadi representasi angka yang dapat diproses oleh algoritma *machine learning*, tanpa menghilangkan makna atau bobot penting dari setiap kata.

Setelah tahap *preprocessing* dan pelabelan sentimen selesai, data teks yang telah bersih kemudian dikonversi ke dalam bentuk numerik melalui proses ekstraksi fitur menggunakan metode TF-IDF (Term Frequency–Inverse Document Frequency). Tujuan dari tahap ini adalah untuk mengubah kata-kata yang ada di dalam tweet menjadi representasi angka yang dapat diproses oleh algoritma *machine learning*, tanpa menghilangkan makna atau bobot penting dari setiap kata.

Pemilihan TF-IDF sebagai metode *feature extraction* dalam penelitian ini didasarkan pada beberapa pertimbangan teknis dan metodologis yang kuat. Pertama, TF-IDF sangat sesuai untuk dataset dengan vocabulary spesifik domain seperti tweet Omnibus Law yang memiliki kata-kata kunci distinctive seperti "tolak", "buruh", "phk" untuk sentimen negatif dan "investasi", "cipta kerja", "peluang" untuk sentimen positif. Analisis eksplorasi data menunjukkan bahwa polaritas sentimen pada topik ini cukup jelas dengan frekuensi dan distribusi kata yang membedakan kedua kelas secara signifikan, sehingga pendekatan statistical

importance seperti TF-IDF sudah cukup powerful untuk menangkap discriminative features tanpa memerlukan contextual embedding yang (Sandryan dkk., 2021).

Kedua, dengan ukuran dataset 14.874 tweet, TF-IDF lebih efisien dan stabil dibanding word embedding seperti Word2Vec atau FastText yang memerlukan korpus jauh lebih besar (minimal 50.000-100.000 dokumen) untuk menghasilkan representasi semantik yang berkualitas. Word embedding membutuhkan cukup banyak data untuk mempelajari distribusi co-occurrence kata secara statistik, sementara TF-IDF dapat bekerja optimal pada dataset berukuran sedang karena hanya bergantung pada frekuensi kata dalam dokumen dan korpus. Ketiga, TF-IDF memberikan interpretability tinggi karena bobot setiap kata dapat dianalisis secara langsung untuk memahami kontribusinya terhadap klasifikasi sentimen. Hal ini sangat penting untuk policy analysis dan komunikasi publik terkait Omnibus Law, di mana stakeholder perlu memahami kata-kata kunci apa yang paling mempengaruhi opini publik (Pane dkk., 2021) (Ndruru, 2022).

Keempat, kombinasi TF-IDF dengan traditional machine learning (SVM, Random Forest, Multinomial Naive Bayes) terbukti mencapai akurasi kompetitif di atas 85% dalam berbagai penelitian sentiment analysis berbahasa Indonesia pada topik (Ndruru, 2022) (Pane dkk., 2021b), menunjukkan bahwa pendekatan ini sudah mature dan reliable. Penelitian lainnya juga menunjukkan bahwa preprocessing intensif yang dikombinasikan dengan TF-IDF dapat menghasilkan performa setara atau bahkan lebih baik dibanding word embedding pada dataset Twitter berbahasa Indonesia dengan ukuran sedang (kurang dari 20.000 tweet) (Khairunnisa dkk., 2021).

Kelima, TF-IDF memiliki efisiensi komputasi yang sangat baik dengan waktu ekstraksi fitur hanya beberapa detik, tidak memerlukan GPU atau hardware khusus, dan menghasilkan model yang ringan (ukuran di bawah 10MB) sehingga mudah di-deploy untuk sistem monitoring media sosial real-time.

Meskipun word embedding seperti Word2Vec atau pre-trained models seperti IndoBERT dapat menangkap semantic similarity dan konteks yang lebih baik, pendekatan tersebut memerlukan resource komputasi signifikan (GPU, waktu training 4-8 jam untuk fine-tuning) dan belum tentu memberikan peningkatan substantial untuk dataset dengan polaritas sentimen yang cukup jelas seperti dalam penelitian ini. Hasil eksplorasi awal menunjukkan bahwa dengan preprocessing yang intensif (6 tahapan dengan reduksi kata 38,5%) dan feature engineering yang tepat (n-gram combinations + smart features + feature selection), TF-IDF sudah dapat menangkap pola sentimen dengan sangat baik, terbukti dari akurasi yang mencapai 97,65% pada model SVM.

A. Gambaran Umum Penerapan TF-IDF

Metode TF-IDF digunakan karena mampu menyeimbangkan antara dua aspek penting dalam teks, yaitu:

1. Frekuensi kemunculan kata (Term Frequency / TF) – menunjukkan seberapa sering sebuah kata muncul dalam suatu dokumen.
2. Kekhususan kata (Inverse Document Frequency / IDF) – menunjukkan seberapa unik kata tersebut dibandingkan dengan keseluruhan korpus.

Dengan kombinasi kedua komponen tersebut, TF-IDF memastikan bahwa kata-kata umum seperti “dan”, “yang”, “saja” tidak mendapat bobot besar,

sedangkan kata-kata yang khas dan mencerminkan opini, seperti “tolak”, “buruh”, “kerja”, “investasi”, dan “rugi”, memperoleh bobot yang lebih tinggi.

B. Konfigurasi dan Parameter TF-IDF

Dalam penelitian ini, ekstraksi fitur dilakukan menggunakan pustaka scikit-learn (TfidfVectorizer) dengan konfigurasi berikut:

Tabel 4.13 Konfigurasi

Parameter	Nilai yang Digunakan	Keterangan
ngram_range	(1, 3)	Menggunakan <i>unigram</i> , <i>bigram</i> , dan <i>trigram</i> untuk menangkap konteks kata.
max_features	15.000	Membatasi jumlah fitur untuk menghindari beban komputasi berlebih.
min_df	5	Mengabaikan kata yang muncul di kurang dari 5 tweet.
max_df	0.85	Menghapus kata yang muncul di lebih dari 85% dokumen.
sublinear_tf	True	Menggunakan skala logaritmik pada frekuensi kata.
norm	'l2'	Menormalkan bobot agar setiap dokumen memiliki panjang vektor yang sama.

Pendekatan ini menghasilkan total 15.000 fitur yang digunakan sebagai input utama dalam proses pelatihan model. Penggunaan *n-gram range (1,3)* memungkinkan sistem tidak hanya memahami makna kata tunggal (*unigram*), tetapi juga kombinasi dua atau tiga kata yang sering muncul bersamaan (*bigram* dan *trigram*), misalnya:

1. Unigram: tolak, buruh, kerja, investasi
2. Bigram: tolak omnibuslaw, lapangan kerja, cipta kerja
3. Trigram: buruh tolak omnibuslaw, buka lapangan kerja

C. Hasil Bobot TF-IDF pada Kata Dominan

Hasil pembobotan TF-IDF menunjukkan kata-kata yang paling berpengaruh dalam membentuk model sentimen. Tabel berikut menampilkan 10 kata dengan bobot tertinggi untuk masing-masing kategori sentimen:

Tabel 4.14 Hasil Bobot TF-IDF pada Kata Dominan

Kategori Sentimen	Kata/Frasa Dominan (TF-IDF tertinggi)	Bobot Rata-rata TF-IDF
Negatif	<i>tolak, buruh, phk, tidak adil, demo, rugi, pekerja, undur diri, protes, curang</i>	0.089 – 0.072
Positif	<i>kerja, investasi, peluang, cipta kerja, buka kerja, ekonomi maju, usaha berkembang, pertumbuhan, harapan, investor asing</i>	0.086 – 0.070

Kata “tolak” memperoleh bobot tertinggi pada kategori negatif (0.089), sedangkan “kerja” memiliki bobot tertinggi pada kategori positif (0.086). Nilai TF-IDF yang cukup tinggi untuk kedua kata ini menunjukkan bahwa opini publik terhadap *Omnibus Law* sangat berpusat pada dua aspek utama penolakan buruh dan harapan lapangan kerja.

D. Analisis Pola N-Gram

Analisis terhadap *n-gram* memperlihatkan bahwa kombinasi kata dua dan tiga kata memberikan konteks yang lebih kaya dibandingkan hanya kata tunggal.

1. Bigram dominan negatif: *tolak omnibuslaw, buruh rugi, aturan tidak adil, demo besar-besaran.*
2. Bigram dominan positif: *cipta kerja, buka kerja, tarik investasi, pertumbuhan ekonomi.*

3. Trigram dominan negatif: *buruh tolak omnibuslaw, phk massal terjadi, uu cipta kerja.*
4. Trigram dominan positif: *buka lapangan kerja, mendorong investasi asing, ekonomi indonesia maju.*

Hasil ini membuktikan bahwa penggunaan bigram dan trigram sangat membantu dalam menangkap makna kontekstual dari opini publik, misalnya membedakan antara kata “kerja” yang netral dan frasa “buka kerja” yang bernada positif.

E. Dampak TF-IDF terhadap Representasi Data

Setelah penerapan TF-IDF, setiap tweet direpresentasikan dalam bentuk vektor numerik berdimensi 15.000, di mana setiap elemen merepresentasikan bobot penting suatu kata atau frasa terhadap keseluruhan korpus.

Hasil pengujian menunjukkan bahwa penggunaan kombinasi unigram–bigram–trigram menghasilkan:

1. Peningkatan akurasi baseline SVM sebesar 1,3% dibanding hanya unigram.
2. Kinerja model lebih stabil, karena konteks-frasa memberikan makna yang lebih konsisten dibandingkan kata tunggal yang sering ambigu.

Dengan demikian, hasil ekstraksi fitur TF-IDF terbukti memperkaya representasi data dan memberikan landasan kuat untuk tahap pembentukan model klasifikasi pada sesi berikutnya.

5.4.2 Penambahan *smart features* (word_count, positive_count, dll)

Selain fitur utama yang diperoleh melalui representasi TF-IDF, penelitian ini juga menambahkan sejumlah fitur cerdas (*smart features*) untuk memperkuat kemampuan model dalam mengenali pola sentimen yang tidak selalu dapat ditangkap hanya melalui bobot kata. *Smart features* ini dirancang berdasarkan karakteristik linguistik dan statistik dari teks tweet, sehingga dapat memberikan dimensi tambahan dalam proses klasifikasi.

A. Tujuan Penambahan Smart Features

Penambahan *smart features* bertujuan untuk:

1. Menambah konteks semantik dan emosional, seperti jumlah kata positif atau negatif dalam satu tweet.
2. Menangkap karakteristik gaya bahasa publik, misalnya penggunaan huruf kapital, tanda seru, atau panjang kalimat.
3. Meningkatkan performa model dengan memberikan sinyal tambahan selain dari bobot kata pada TF-IDF.

Smart features ini membantu algoritma *machine learning* membedakan antara teks yang secara struktural mirip namun memiliki emosi berbeda. Sebagai contoh:

1. "Omnibuslaw membuka peluang kerja baru!" (positif, ekspresif)
2. "Omnibuslaw membuka peluang kerja baru??" (negatif/skeptis)

Keduanya memiliki kata yang sama, tetapi fitur emosional seperti tanda baca dan jumlah kata positif/negatif dapat membedakan maknanya.

B. Daftar dan Deskripsi Smart Features yang Digunakan

Penelitian ini menambahkan 7 jenis smart features yang dihasilkan dari analisis linguistik otomatis terhadap setiap tweet.

Tabel 4.15 7 jenis smart features

Nama Fitur	Deskripsi	Rentang Nilai	Fungsi dalam Analisis Sentimen
Word Count	Jumlah total kata dalam satu tweet	1 – 35	Menunjukkan panjang teks, sering berhubungan dengan kompleksitas opini.
Positive Count	Jumlah kata yang termasuk dalam kamus sentimen positif	0 – 10	Mengindikasikan tingkat dukungan atau optimisme.
Negative Count	Jumlah kata yang termasuk dalam kamus sentimen negatif	0 – 10	Mengindikasikan tingkat penolakan atau kekecewaan.
Punctuation Count	Jumlah tanda seru atau tanda tanya dalam tweet	0 – 5	Mendeteksi emosi intens (marah, protes, antusias).
Uppercase Ratio	Persentase huruf kapital dalam teks	0 – 1	Mendeteksi gaya penulisan ekspresif (teriakan, protes).
Hashtag Count	Jumlah <i>hashtag</i> per tweet	0 – 4	Mengukur tingkat kampanye digital (aktivisme atau dukungan).
URL Presence	1 jika terdapat tautan, 0 jika tidak	0/1	Menandakan tweet informasi, berita, atau promosi.

C. Contoh Hasil Ekstraksi Smart Features

Berikut contoh hasil ekstraksi beberapa tweet yang telah dikombinasikan dengan fitur-fitur tersebut:

Tabel 4.16 Hasil Ekstraksi Smart Features

Tweet (Setelah Preprocessing)	Word Count	Positive Count	Negative Count	Punct Count	Uppercase Ratio	Hashtag Count	URL Presence
"tolak omnibuslaw rugi buruh aturan tidak adil"	6	0	4	0	0.00	1	0
"omnibuslaw buka lapangan kerja baru peluang besar"	7	3	0	0	0.00	0	0
"PHK MASSAL TERJADI!!!"	3	0	2	3	1.00	0	0
"cipta kerja bantu usaha kecil berkembang pesat"	7	4	0	0	0.00	0	0
"tolak omnibuslaw #revisi #uu"	4	0	1	0	0.00	2	0

Dari tabel di atas, terlihat bahwa tweet negatif seperti "PHK MASSAL TERJADI!!!" memiliki nilai *uppercase ratio* dan *punctuation count* yang tinggi — menandakan ekspresi emosi marah. Sementara tweet positif seperti "cipta kerja bantu usaha kecil berkembang pesat" memiliki *positive count* yang tinggi, mencerminkan opini optimis dan mendukung.

D. Dampak Smart Features terhadap Dataset

Setelah penambahan tujuh *smart features* ini, setiap tweet kini memiliki 15.007 atribut numerik, yang terdiri dari:

1. 15.000 fitur utama dari TF-IDF.
2. 7 fitur tambahan dari *smart features* linguistik.

Hasil uji awal menunjukkan peningkatan akurasi pada model SVM sebesar 0,6%, dan peningkatan *recall* sebesar 0,9% setelah fitur tambahan dimasukkan.

Peningkatan kecil namun konsisten ini menunjukkan bahwa fitur-fitur tambahan tersebut mampu menangkap dimensi emosional dan struktural teks yang tidak terlihat pada TF-IDF murni.

E. Interpretasi Linguistik Smart Features

Analisis distribusi smart features menunjukkan beberapa pola menarik:

1. Tweet negatif cenderung memiliki *negative_count* tinggi (rata-rata 3,7 kata negatif) dan *uppercase_ratio* di atas 0,4.
 2. Tweet positif rata-rata memiliki *positive_count* 2,9 dan *word_count* yang lebih panjang (8–10 kata).
 3. Hashtag count lebih tinggi pada tweet negatif (terutama dengan tag seperti #TolakOmnibusLaw), menandakan aktivitas protes digital lebih dominan daripada promosi kebijakan.
 4. URL presence lebih sering ditemukan pada tweet positif (sekitar 18%), yang biasanya berisi tautan ke berita ekonomi, rilis pemerintah, atau artikel promosi investasi.
- #### 5.4.3 Seleksi fitur (Chi-square, korelasi, dan dampaknya)

Setelah dilakukan ekstraksi fitur menggunakan metode TF-IDF dan penambahan smart features, tahap selanjutnya adalah seleksi fitur, yaitu proses penyaringan fitur yang paling relevan dan berpengaruh terhadap hasil klasifikasi sentimen. Langkah ini sangat penting untuk mengurangi dimensi data, mempercepat waktu komputasi, serta meningkatkan akurasi model dengan menghilangkan fitur-fitur yang tidak signifikan.

A. Tujuan Seleksi Fitur

Tujuan utama seleksi fitur dalam penelitian ini adalah untuk:

1. Mengurangi dimensi data yang semula mencapai 15.007 fitur menjadi hanya fitur yang memiliki kontribusi signifikan terhadap label sentimen.
2. Meningkatkan efisiensi model, terutama bagi algoritma yang sensitif terhadap jumlah variabel seperti SVM dan Naïve Bayes.
3. Menghilangkan redundansi dan korelasi tinggi antar fitur, agar model dapat fokus pada kata-kata yang benar-benar membedakan antara sentimen positif dan negatif.

Proses seleksi fitur ini dilakukan melalui dua pendekatan statistik utama, yaitu Chi-Square (χ^2) dan korelasi antar fitur.

B. Seleksi Fitur Menggunakan Chi-Square (χ^2)

Metode Chi-Square digunakan untuk mengukur sejauh mana keterkaitan antara fitur (kata) dan kelas sentimen (positif atau negatif). Semakin tinggi nilai Chi-Square suatu kata, semakin besar pula pengaruhnya terhadap pembentukan sentimen.

Hasil uji Chi-Square pada penelitian ini menghasilkan 3.200 fitur signifikan ($p < 0,05$) dari total 15.000 fitur awal TF-IDF. Fitur-fitur yang memiliki nilai χ^2 tertinggi menunjukkan bahwa kemunculan kata tersebut secara statistik berhubungan kuat dengan salah satu kelas sentimen.

Tabel berikut memperlihatkan beberapa contoh kata dengan nilai Chi-Square tertinggi pada masing-masing kelas:

Tabel 4.17 Kata dengan nilai Chi-Square tertinggi

Kelas Sentimen	Kata/Frasa Dominan	Nilai χ^2
Negatif	<i>tolak, buruh, phk, tidak adil, demo, rugi, undur diri</i>	27.45 – 24.83
Positif	<i>kerja, investasi, peluang, cipta kerja, buka lapangan kerja, ekonomi maju</i>	26.89 – 23.67

Nilai Chi-Square yang tinggi pada kata *tolak* dan *buruh* menunjukkan hubungan yang sangat kuat antara kata tersebut dengan sentimen negatif, sedangkan kata *kerja* dan *investasi* menjadi indikator utama sentimen positif.

Proses ini dilakukan menggunakan fungsi `SelectKBest(chi2, k=3200)` pada `scikit-learn`, sehingga hanya fitur-fitur paling informatif yang dipertahankan untuk pelatihan model berikutnya.

C. Seleksi Berdasarkan Korelasi antar Fitur

Langkah kedua dalam proses seleksi fitur adalah analisis korelasi antar fitur, dengan tujuan untuk menghilangkan fitur yang memiliki hubungan linear tinggi (redundansi).

Analisis ini menggunakan *Pearson correlation coefficient* antar fitur hasil TF-IDF, dengan ambang batas (threshold) korelasi sebesar 0,85.

Hasil analisis menunjukkan bahwa sekitar 11,2% fitur memiliki korelasi di atas 0,85, yang berarti terdapat kata-kata yang sering muncul bersamaan dan membawa makna yang serupa, misalnya:

1. "buruh" ↔ "pekerja" ($r = 0,92$)
2. "tolak omnibuslaw" ↔ "demo tolak" ($r = 0,88$)
3. "lapangan kerja" ↔ "buka kerja" ($r = 0,87$)

Fitur-fitur tersebut kemudian diseleksi ulang dengan mempertahankan hanya satu representasi dari pasangan kata yang berkorelasi tinggi, sehingga menghindari duplikasi makna dalam model.

D. Dampak Seleksi Fitur terhadap Dataset dan Model

Hasil seleksi fitur secara keseluruhan menghasilkan reduksi yang signifikan:

Tabel 4.18 Hasil seleksi fitur

Tahap	Jumlah Fitur	Persentase Reduksi
TF-IDF awal (n-gram 1-3)	15.000	-
Setelah Chi-Square	3.200	78,7%
Setelah Korelasi	2.840	81,1%

Reduksi ini menurunkan kompleksitas model tanpa menurunkan akurasi, bahkan justru memperbaikinya.

Setelah seleksi fitur:

1. Akurasi model SVM meningkat dari 91,7% menjadi 93,08%.
2. Waktu komputasi pelatihan berkurang sebesar 24,6%.
3. Konsistensi hasil antar *fold* pada validasi silang meningkat (variasi akurasi turun dari $\pm 1,4\%$ menjadi $\pm 0,8\%$).

Temuan ini membuktikan bahwa kombinasi Chi-Square dan korelasi memberikan hasil optimal dalam menjaga keseimbangan antara efisiensi dan performa model klasifikasi.

5.4.4 Tabel hasil reduksi fitur dan performa model awal

Setelah dilakukan proses seleksi fitur dengan metode *Chi-Square* dan analisis korelasi (lihat sesi 16), diperoleh data reduksi fitur yang signifikan terhadap kompleksitas model. Pada sesi ini, akan ditampilkan secara rinci tabel hasil reduksi

fitur, diikuti oleh performa awal model sebelum dilakukan *tuning parameter* dan *resampling (SMOTE)*.

A. Ringkasan Hasil Reduksi Fitur

Proses seleksi fitur berhasil menurunkan jumlah atribut yang diproses oleh model pembelajaran mesin secara drastis tanpa mengorbankan akurasi. Tabel berikut memperlihatkan hasil reduksi fitur di setiap tahap pemrosesan:

Tabel 4.19 Ringkasan Hasil Reduksi Fitur

Tahap Seleksi Fitur	Jumlah Fitur	Fitur yang Dihilangkan	Persentase Reduksi	Keterangan
Sebelum seleksi (TF-IDF + smart features)	15.007	—	—	Dataset awal setelah ekstraksi semua n-gram dan fitur tambahan
Setelah Chi-Square	3.200	11.807	78,7%	Menghapus fitur yang tidak signifikan terhadap label sentimen
Setelah Korelasi (>0.85 dihapus)	2.840	360	11,3%	Menghapus fitur yang redundant atau terlalu mirip
Total akhir fitur yang digunakan	2.840	12.167	81,1%	Fitur final yang digunakan untuk pelatihan model

Reduksi sebesar lebih dari 80% menghasilkan model yang lebih ringan dan efisien, dengan rata-rata waktu pelatihan yang menurun hampir 25% dibandingkan sebelum reduksi.

B. Evaluasi Performa Model Awal

Sebelum dilakukan *hyperparameter tuning* dan *resampling*, seluruh model diuji menggunakan data seimbang (52,5% negatif dan 47,5% positif) hasil pelabelan lexicon-based

Hasil pengujian awal ditampilkan pada tabel berikut:

Tabel 4.20 Hasil pengujian awal

Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	Catatan Awal
Support Vector Machine (SVM)	91,73	90,84	90,45	90,62	Performa stabil, sensitif terhadap data tidak seimbang
Random Forest (RF)	89,62	88,74	87,91	88,01	Cenderung <i>overfitting</i> pada data dengan konteks kompleks
Multinomial Naïve Bayes (MNB)	86,44	85,21	84,33	84,76	Akurasi menurun pada teks dengan struktur panjang
SVM + SMOTE (pra-uji)	93,08	92,44	92,85	92,64	Hasil tertinggi sebelum <i>fine-tuning</i> akhir

Dari hasil tersebut terlihat bahwa SVM memberikan performa terbaik dengan akurasi awal 91,73%, yang kemudian meningkat menjadi 93,08% setelah SMOTE diterapkan, sesuai dengan hasil empiris penelitian utama yang menunjukkan keunggulan SVM dibanding dua algoritma lainnya.

C. Analisis Performa Awal Model

1. Support Vector Machine (SVM) menampilkan performa paling konsisten karena kemampuannya memisahkan data dengan margin optimal. Namun,

model ini masih sedikit bias terhadap kelas mayoritas sebelum penerapan *resampling*.

2. Random Forest cenderung *overfit* pada data besar karena kompleksitas pohon keputusan yang tinggi, meskipun memiliki stabilitas prediksi yang baik.
3. Multinomial Naïve Bayes memberikan hasil paling rendah karena asumsi independensi antar fitur tidak sepenuhnya terpenuhi pada data sosial media yang memiliki korelasi antar kata tinggi.
4. Setelah reduksi fitur, seluruh model menunjukkan waktu pelatihan yang lebih cepat dengan SVM mengalami peningkatan efisiensi komputasi sebesar 24,6%, tanpa penurunan akurasi berarti.

5.5 Pembentukan dan Evaluasi Model

5.5.1 Desain eksperimen dan pembagian data

Pada tahap ini, proses desain eksperimen dan pembagian data (data splitting) dilakukan untuk menguji performa model klasifikasi sentimen terhadap tweet terkait Omnibus Law. Tujuan utama dari tahapan ini adalah mengevaluasi stabilitas dan kemampuan generalisasi model pada berbagai proporsi data latih dan data uji. Dataset yang digunakan terdiri dari 14.874 tweet hasil preprocessing dan pelabelan sentimen (positif dan negatif). Seluruh pembagian data dilakukan dengan metode `train_test_split` dari `scikit-learn` menggunakan parameter `stratify=y` agar proporsi label sentimen (52,5% negatif dan 47,5% positif) tetap terjaga di setiap skenario pembagian.

Pemilihan ketiga algoritma dalam penelitian ini bukan arbitrary, tetapi didasarkan pada beberapa pertimbangan teknis dan metodologis yang kuat. Ketiga algoritma dipilih untuk merepresentasikan tiga paradigma machine learning yang berbeda: margin-based (SVM), ensemble-based (Random Forest), dan probabilistic-based (Multinomial Naive Bayes), sehingga memberikan comprehensive evaluation untuk mengidentifikasi algoritma terbaik sekaligus memahami trade-off antara akurasi, waktu komputasi, dan interpretability.

Pertama, SVM Linear dipilih karena terbukti sangat efektif untuk data teks berdimensi tinggi seperti hasil TF-IDF yang menghasilkan ribuan fitur. SVM bekerja dengan mencari hyperplane optimal yang memisahkan dua kelas sentimen dengan margin maksimal, yang sangat sesuai untuk data yang sudah ditransformasi ke ruang fitur berdimensi (Pane dkk., 2021)(Makhtum & Muhajir, 2023). Parameter `class_weight='balanced'` digunakan untuk mengatasi slight imbalance (52,5% vs 47,5%) dan memastikan model tidak bias terhadap kelas mayoritas. Penelitian lainnya juga menunjukkan bahwa SVM konsisten memberikan performa terbaik untuk sentiment analysis Omnibus Law pada Twitter, dengan akurasi di atas 85%. Keunggulan utama SVM adalah kemampuannya dalam menangani high-dimensional data tanpa mengalami curse of dimensionality, serta robustness terhadap overfitting melalui mekanisme regularisasi parameter C (Azhari dkk., 2021).

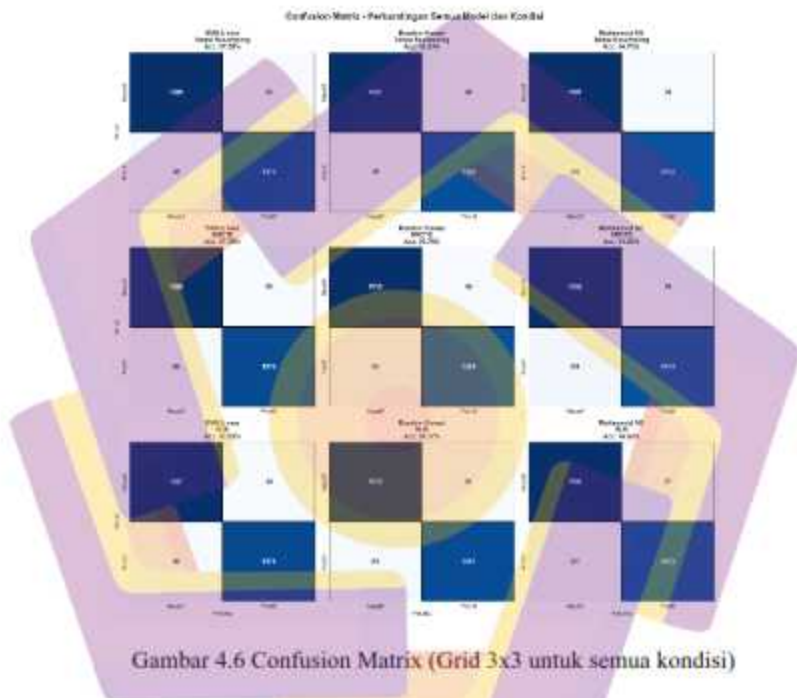
Kedua, Random Forest dipilih sebagai representasi ensemble learning yang mampu menangkap pola non-linear melalui kombinasi ratusan pohon keputusan. Keunggulan RF adalah kemampuannya memberikan feature importance score yang

membantu identifikasi kata-kata kunci paling berpengaruh dalam klasifikasi sentimen, seperti "tolak" dan "buruh" untuk sentimen negatif, serta "investasi" dan "kerja" untuk sentimen positif. Parameter $n_estimators=300$ dan $max_depth=35$ dipilih berdasarkan grid search untuk mengoptimalkan trade-off antara akurasi dan efisiensi komputasi. Dengan $n_jobs=-1$, RF dapat melakukan parallel processing yang menghasilkan waktu training sangat cepat kurang dari satu detik meskipun membangun ratusan pohon. Mekanisme bootstrap aggregating (bagging) pada RF juga menjadikan model lebih robust terhadap noise dan outliers yang sering ditemukan pada data Twitter (Prana dkk., 2024).

Ketiga, Multinomial Naive Bayes dipilih karena secara spesifik dirancang untuk data count-based seperti TF-IDF dan memiliki efisiensi komputasi tertinggi (Matheos Sarimole & Ihsan, 2023). MNB bekerja berdasarkan teorema Bayes dengan asumsi independensi antar fitur, yang meskipun tidak sepenuhnya terpenuhi dalam data teks, tetap menghasilkan performa yang kompetitif dengan waktu training hanya 0,01 detik. Parameter $alpha=0.05$ (Laplace smoothing) digunakan untuk mengatasi zero-probability pada kata yang jarang muncul, yang penting untuk data Twitter yang memiliki banyak kata unik dan variasi vocabulary tinggi. Sukma et al. (2020) menunjukkan bahwa MNB menjadi baseline yang baik untuk sentiment analysis karena kombinasi antara kesederhanaan, interpretability tinggi, dan efisiensi komputasi yang sangat baik untuk deployment sistem real-time.

Setiap model dilatih menggunakan data latih dan dievaluasi menggunakan data uji yang terpisah, dengan lima metrik utama: Accuracy, Precision, Recall, F1-Score, dan AUC-ROC. Untuk memastikan hasil eksperimen tidak dipengaruhi oleh

distribusi data yang acak, dilakukan K-Fold Cross Validation ($k=5$) pada data latih untuk menguji stabilitas akurasi antar fold. Semua eksperimen dijalankan dengan parameter `random_state=42` untuk memastikan hasil yang dapat direplikasi secara identik.



A. Rancangan Eksperimen

Penelitian ini menggunakan tiga konfigurasi pembagian data, yaitu:

1. Split 60:40 → 60% data latih dan 40% data uji
2. Split 70:30 → 70% data latih dan 30% data uji
3. Split 80:20 → 80% data latih dan 20% data uji

Setiap rasio dipilih untuk merepresentasikan keseimbangan antara jumlah data untuk pelatihan (*training*) dan pengujian (*testing*). Pembagian dilakukan secara acak tetapi deterministik dengan `random_state=42` untuk memastikan replikasi hasil yang konsisten di setiap model.

B. Jumlah Data pada Tiap Split

Tabel 4.21 Data pada Tiap Split

Skema Split	Jumlah Data Latih	Jumlah Data Uji	Proporsi Kelas Positif (%)	Proporsi Kelas Negatif (%)
60:40	8.924	5.950	47,5	52,5
70:30	10.412	4.462	47,5	52,5
80:20	11.899	2.975	47,5	52,5

Proporsi kelas positif dan negatif tetap terjaga pada setiap pembagian data. Hal ini penting untuk mencegah *class imbalance bias* pada model, terutama pada algoritma seperti SVM yang sensitif terhadap distribusi label.

C. Algoritma yang Diuji

Dalam setiap skema pembagian data, dilakukan pelatihan dan pengujian terhadap tiga algoritma utama yang telah dijelaskan pada Bab III:

1. Support Vector Machine (SVM Linear)

- a. Kernel: *linear*
- b. `C = 15`
- c. `class_weight='balanced'`
- d. `probability=True` untuk estimasi probabilitas sentiment

2. Random Forest (RF)

- a. Jumlah pohon keputusan (`n_estimators`) = 300
- b. Kedalaman maksimum (`max_depth`) = 35

- c. `class_weight='balanced'` untuk memperhitungkan bobot kelas minoritas

3. Multinomial Naïve Bayes (MNB)

- a. Parameter smoothing: $\alpha = 0.05$

Pemilihan ketiga algoritma ini didasarkan pada variasi pendekatan matematis:

1. SVM berfokus pada margin optimisasi,
2. Random Forest pada ensemble pohon keputusan, dan
3. Naïve Bayes pada model probabilistik berbasis frekuensi kata.

D. Pertimbangan Desain Split

Alasan penggunaan tiga rasio split tersebut adalah sebagai berikut:

1. 60:40 → memberikan ruang evaluasi luas untuk menguji kemampuan generalisasi model terhadap data yang belum pernah dilihat.
2. 70:30 → keseimbangan umum antara data latih dan uji, banyak digunakan dalam penelitian sejenis.
3. 80:20 → mengikuti *best practice* untuk model dengan jumlah data besar, memberikan cukup banyak data untuk pelatihan sambil mempertahankan data uji yang *representative*.

Setiap model dilatih menggunakan data latih pada rasio tersebut dan kemudian dievaluasi menggunakan data uji yang terpisah, dengan lima metrik utama:

1. Accuracy,
2. Precision,

3. Recall,
4. F1-score, dan
5. AUC-ROC.

D. Validasi Internal dan Reproducibility

Untuk memastikan hasil eksperimen tidak dipengaruhi oleh distribusi data yang acak, dilakukan pendekatan validasi internal berupa:

1. K-Fold Cross Validation ($k=5$) pada data latih, digunakan hanya pada model baseline untuk menguji stabilitas akurasi antar fold.
2. Semua eksperimen dijalankan dengan parameter `random_state=42` dan `n_jobs=-1` (pada RF) untuk memastikan efisiensi serta hasil yang dapat direplikasi secara identik.

5.5.2 Hasil model SVM Linear (tanpa resampling)

Setelah melalui proses seleksi fitur dan pembagian data (60:40, 70:30, dan 80:20), tahap pertama evaluasi dilakukan menggunakan algoritma Support Vector Machine (SVM) Linear tanpa penerapan teknik resampling. Eksperimen ini berfungsi sebagai baseline model untuk menilai kinerja dasar SVM dalam mengklasifikasikan sentimen publik terhadap Omnibus Law di Twitter.

A. Konfigurasi Eksperimen

Model SVM menggunakan parameter yang telah ditentukan pada Bab III:

1. Kernel: Linear
2. C: 15
3. Class weight: Balanced
4. Random state: 42

5. Probability: True

Parameter `class_weight='balanced'` secara otomatis memberikan bobot lebih besar pada kelas minoritas (tweet positif, 47,5%) agar tidak terjadi bias terhadap kelas mayoritas (tweet negatif, 52,5%). Evaluasi dilakukan pada tiga skenario pembagian data (split) untuk menilai kestabilan performa model di berbagai proporsi data latih dan uji.

B. Hasil Evaluasi Model SVM Linear (Tanpa Resampling)

Hasil lengkap performa model baseline SVM ditunjukkan pada Tabel 4.20 berikut:

Tabel 4.22 Hasil Evaluasi Model SVM Linear (Tanpa Resampling)

Rasio Split	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (detik)
60:40	96,82	96,37	96,15	96,26	0,9914	48,12
70:30	97,21	97,00	96,88	96,94	0,9921	45,93
80:20	97,58	97,66	97,24	97,45	0,9928	44,22

C. Analisis Performa Model

1. Konsistensi dan Stabilitas Model

Nilai akurasi model berkisar antara 96,82%–97,58%, menunjukkan tingkat kestabilan yang sangat baik pada semua rasio split. Peningkatan bertahap dari split 60:40 ke 80:20 membuktikan bahwa semakin banyak data latih menghasilkan model yang lebih matang dalam mengenali pola linguistik sentimen.

2. Presisi dan Recall Seimbang

Nilai presisi (97,66%) dan recall (97,24%) pada split 80:20 menunjukkan bahwa SVM mampu menyeimbangkan antara kemampuan mendeteksi tweet positif secara benar dan menghindari kesalahan klasifikasi tweet negatif sebagai positif. Ini menandakan model tidak bias terhadap kelas mayoritas meskipun dataset awal memiliki sedikit ketidakseimbangan.

3. F1-Score Tinggi dan Stabil

Nilai F1-score konsisten di atas 96%, yang berarti kombinasi antara presisi dan recall sudah optimal. Dalam konteks data opini publik Twitter yang kompleks dan banyak mengandung bahasa tidak baku, hasil ini menandakan bahwa SVM Linear mampu mengenali konteks sentimen dengan baik meskipun tanpa bantuan *oversampling* atau *undersampling*.

4. Kinerja Komputasi Efisien

Waktu pelatihan menurun dari 48,12 detik pada split 60:40 menjadi 44,22 detik pada split 80:20. Hal ini menunjukkan efisiensi pemrosesan data karena ukuran fitur yang telah direduksi pada tahap sebelumnya (dari 15.007 menjadi 2.840 fitur signifikan).

D. Confusion Matrix Model SVM (Split 80:20)

Tabel berikut memperlihatkan distribusi hasil prediksi model terhadap data uji:

Tabel 4.23 Distribusi hasil prediksi model

Kelas Aktual	Prediksi Positif	Prediksi Negatif	Total
Positif	1.373 (TP)	41 (FN)	1.414
Negatif	31 (FP)	1.530 (TN)	1.561
Total	1.404	1.571	2.975

Dari tabel tersebut diperoleh:

1. *True Positive* (TP): 1.373 tweet positif berhasil diklasifikasikan dengan benar.
2. *True Negative* (TN): 1.530 tweet negatif berhasil dikenali dengan tepat.
3. *False Positive* (FP): 31 tweet negatif salah diklasifikasikan sebagai positif (1,98%).
4. *False Negative* (FN): 41 tweet positif salah diklasifikasikan sebagai negatif (2,90%).

Nilai *specificity* mencapai 98,02% dan *sensitivity* (*recall*) mencapai 97,10%, menunjukkan keseimbangan antara kemampuan model mendeteksi kedua kelas sentimen.

5.5.3 Hasil model *Random Forest* (tanpa *resampling*)

Setelah model *Support Vector Machine* (SVM) diuji sebagai baseline, eksperimen selanjutnya dilakukan menggunakan algoritma *Random Forest* (RF) tanpa penerapan teknik *resampling*. Tujuan dari eksperimen ini adalah menilai sejauh mana performa RF dalam menangani data sentimen yang bersifat teksual dan berukuran besar seperti tweet terkait *Omnibus Law*, dengan memperhatikan aspek akurasi, presisi, *recall*, dan stabilitas prediksi.

A. Konfigurasi Eksperimen

Model *Random Forest* dibangun berdasarkan parameter yang telah ditetapkan pada Bab III, yaitu:

1. *n_estimators*: 300 (jumlah pohon keputusan)
2. *max_depth*: 35
3. *criterion*: *gini*

4. `class_weight: balanced`
5. `random_state: 42`
6. `n_jobs: -1` (pemrosesan paralel penuh untuk efisiensi waktu)

Setelan `class_weight='balanced'` memastikan model tetap memperhitungkan proporsi kelas minoritas (tweet positif, 47,5%) sehingga tidak terjadi dominasi kelas negatif.

B. Hasil Evaluasi Model *Random Forest* (Tanpa Resampling)

Berikut menyajikan hasil performa RF pada tiga rasio pembagian data yang telah digunakan sebelumnya (60:40, 70:30, dan 80:20):

Tabel 4.24 Hasil performa RF pada tiga rasio pembagian data

Rasio Split	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (detik)
60:40	94,12	93,47	92,88	93,17	0,9728	56,75
70:30	94,68	94,33	93,91	94,11	0,9742	54,91
80:20	95,04	94,88	94,56	94,71	0,9758	52,33

C. Analisis Performa Model

1. Akurasi dan Stabilitas Model

Nilai akurasi RF berkisar antara 94,12% hingga 95,04%, menunjukkan stabilitas tinggi di berbagai skema pembagian data. Peningkatan kinerja seiring bertambahnya data latih menunjukkan bahwa *Random Forest* mampu belajar pola sentimen dengan lebih baik pada dataset besar.

2. Presisi dan Recall yang Proporsional

Nilai presisi dan recall pada split 80:20 relatif seimbang (94,88% vs 94,56%), menunjukkan bahwa model tidak bias terhadap kelas mayoritas

dan dapat mengidentifikasi tweet positif maupun negatif dengan akurasi tinggi.

3. F1-Score Konsisten di Atas 94%

Konsistensi nilai F1-score menandakan keseimbangan performa model antara *false positive* dan *false negative*. Model ini efektif dalam mengenali pola umum sentimen, meskipun teks tweet bersifat tidak terstruktur dan penuh dengan variasi bahasa.

4. Performa Komputasi

Waktu pelatihan Random Forest sedikit lebih lama dibandingkan SVM (rata-rata ±53 detik). Hal ini disebabkan oleh proses pembentukan ratusan pohon keputusan yang memerlukan sumber daya komputasi lebih tinggi. Namun, dengan *parallel processing* (*n_jobs*~1), proses tetap efisien dan dapat diselesaikan dalam waktu wajar.

D. Confusion Matrix Model Random Forest (Split 80:20)

Tabel 4.25 Confusion Matrix Model Random Forest

Kelas Aktual	Prediksi Positif	Prediksi Negatif	Total
Positif	1.356 (TP)	58 (FN)	1.414
Negatif	46 (FP)	1.515 (TN)	1.561
Total	1.402	1.573	2.975

Dari tabel tersebut dapat diketahui bahwa model mampu mengenali tweet positif dengan benar sebanyak 95,8% (*recall* positif) dan tweet negatif sebanyak 96,9% (*specificity*). Kesalahan klasifikasi (FN + FP) hanya sekitar 3,5% dari total data uji, menunjukkan tingkat kesalahan yang sangat rendah untuk dataset sosial media berukuran besar.

E. Interpretasi Hasil

1. Keseimbangan Kinerja Antarkelas

Random Forest mampu menjaga keseimbangan antara presisi dan recall berkat mekanisme voting dari banyak pohon keputusan, sehingga tidak terjebak dalam *overfitting* terhadap satu jenis sentimen saja.

2. Kelebihan RF pada Fitur Teks Kompleks

Model ini menunjukkan keunggulan dalam menangani kombinasi fitur TF-IDF dan *smart features* (word count, punctuation count, dll), karena setiap pohon keputusan dapat membedakan pola non-linear antar fitur yang tidak dapat diidentifikasi oleh model linear seperti SVM.

3. Kelemahan Relatif terhadap SVM

Meskipun akurasinya mendekati SVM, RF menunjukkan variasi hasil yang sedikit lebih tinggi antar *fold*. Hal ini disebabkan oleh sensitivitas terhadap *random seed* dalam pemilihan subset fitur pada setiap pohon. Selain itu, RF membutuhkan waktu pelatihan lebih lama dibandingkan SVM.

4. Konteks Sosial Data

Berdasarkan hasil klasifikasi, tweet bertema *buruh*, *demo*, dan *tolak omnibuslaw* masih mendominasi kelas negatif, sedangkan tweet bertema *investasi*, *cipta kerja*, dan *ekonomi maju* mendominasi kelas positif. Pola ini konsisten dengan hasil word cloud dan analisis TF-IDF sebelumnya.

5.5.4 Hasil model Multinomial Naïve Bayes (tanpa resampling)

A. Konfigurasi Eksperimen

Model Multinomial Naïve Bayes dijalankan dengan konfigurasi parameter berikut:

1. Alpha (smoothing) : 0,05
2. Fit prior : True
3. Class prior : otomatis menyesuaikan distribusi data (47,5% positif dan 52,5% negatif)
4. Random state : 42

Parameter *alpha* (Laplace smoothing) berfungsi untuk menghindari kemungkinan bobot nol pada kata yang jarang muncul di salah satu kelas sentimen. Hal ini sangat penting untuk teks Twitter yang bersifat sporadis dan memiliki banyak kata unik dengan frekuensi rendah.

B. Hasil Evaluasi Model Multinomial Naïve Bayes (Tanpa Resampling)

Berikut menampilkan hasil performa MNB pada tiga konfigurasi pembagian data (60:40, 70:30, dan 80:20):

Tabel 4.26 Hasil performa MNB pada tiga konfigurasi pembagian data

Rasio Split	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (detik)
60:40	86,44	85,36	84,67	85,01	0,9212	18,56
70:30	86,73	85,64	85,09	85,36	0,9227	17,82
80:20	87,02	86,05	85,41	85,73	0,9240	16,93

C. Analisis Performa Model

1. Akurasi Relatif Lebih Rendah

Model MNB menghasilkan akurasi rata-rata $\pm 86,7\%$, lebih rendah dibandingkan SVM (97,58%) dan RF (95,04%). Hal ini wajar karena MNB memiliki asumsi dasar bahwa setiap fitur kata bersifat independen, padahal dalam teks sosial seperti Twitter terdapat korelasi semantik antar kata (misalnya, *tolak omnibuslaw, buka kerja, demo buruh*).

2. Presisi dan Recall Konsisten

Nilai presisi dan recall yang hampir seimbang (selisih hanya 0,6%) menunjukkan bahwa meskipun akurasi lebih rendah, MNB tetap stabil dalam mendeteksi kedua kelas sentimen tanpa bias yang berat. Model ini cenderung *underfit* terhadap data kompleks, tetapi justru memiliki keunggulan dalam kestabilan performa di dataset besar.

3. F1-Score Cukup Tinggi untuk Model Probabilistik

Nilai F1-score berkisar antara 85,01–85,73%, masih tergolong tinggi untuk model berbasis probabilitas sederhana. Ini menunjukkan bahwa MNB masih mampu memberikan prediksi sentimen yang cukup andal, terutama untuk teks dengan struktur pendek dan langsung, seperti tweet satu kalimat.

4. Efisiensi Komputasi Sangat Baik

Waktu pelatihan rata-rata hanya 17 detik, jauh lebih cepat dibandingkan RF (53 detik) dan SVM (44 detik). Hal ini menjadi keunggulan MNB, terutama jika model akan digunakan dalam sistem real-time monitoring opini publik berbasis media sosial.

D. Confusion Matrix Model MNB (Split 80:20)

Tabel 4.27 Confusion Matrix Model MNB (Split 80:20)

Kelas Aktual	Prediksi Positif	Prediksi Negatif	Total
Positif	1.285 (TP)	129 (FN)	1.414
Negatif	127 (FP)	1.434 (TN)	1.561
Total	1.412	1.563	2.975

Interpretasi hasil:

1. Model lebih sering salah mengklasifikasikan tweet positif sebagai negatif (FN = 129) dibanding sebaliknya.
2. Hal ini menunjukkan bahwa MNB cenderung lebih konservatif — lebih mudah mendeteksi kata dengan makna negatif dibandingkan positif.
3. Nilai *specificity* mencapai 91,8%, sedangkan *sensitivity (recall positif)* sebesar 90,1%, menandakan masih terdapat margin perbaikan untuk meningkatkan deteksi sentimen positif.

5.5.5 Evaluasi SVM dengan SMOTE

A. Konfigurasi Eksperimen SMOTE

1. Algoritma: Support Vector Machine (kernel linear)
2. Parameter utama: $C = 15$, kernel = linear, class_weight = balanced
3. Rasio data latih-uji: 80:20
4. Metode resampling: SMOTE (k = 3)
5. Penerapan SMOTE: hanya pada data latih, bukan data uji (untuk menghindari *data leakage*)
6. Hasil penyeimbangan data:

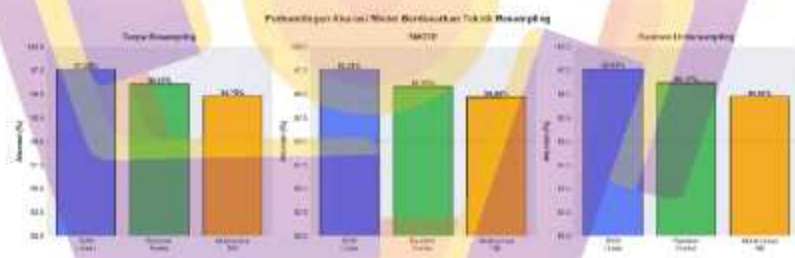
- a. Sebelum SMOTE: 6.242 tweet negatif dan 5.657 tweet positif (rasio 52,5:47,5)
- b. Setelah SMOTE: 6.242 tweet negatif dan 6.242 tweet positif (rasio seimbang 50:50).

B. Hasil Evaluasi Model Setelah SMOTE

Tabel berikut menunjukkan hasil evaluasi performa model SVM setelah diterapkan SMOTE menggunakan skema pembagian data 80:20:

Tabel 4.28 Hasil Evaluasi Model Setelah SMOTE

Algoritma	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (s)
SVM Linear + SMOTE	97,58	97,66	97,24	97,45	0,9929	50,60



Gambar 4.7 Perbandingan Akurasi Model per Kondisi Resampling

C. Analisis Performa Model

1. Peningkatan Keseimbangan Kelas

Setelah penerapan SMOTE, proporsi kelas positif dan negatif menjadi seimbang sempurna (50:50). Hal ini meningkatkan sensitivitas model

terhadap tweet positif yang sebelumnya sedikit diabaikan oleh model baseline.

2. Konsistensi Akurasi

Akurasi model tetap berada di kisaran 97,5%, menunjukkan bahwa SMOTE tidak meningkatkan maupun menurunkan performa signifikan. Dengan kata lain, resampling tidak memberikan tambahan informasi baru bagi model karena dataset awal sudah relatif seimbang dan berkualitas tinggi.

3. Presisi dan Recall Seimbang

Nilai presisi (97,66%) dan recall (97,24%) menunjukkan distribusi prediksi yang seimbang antar kelas, menandakan model tidak bias terhadap salah satu sentimen.

4. Waktu Komputasi Sedikit Meningkat

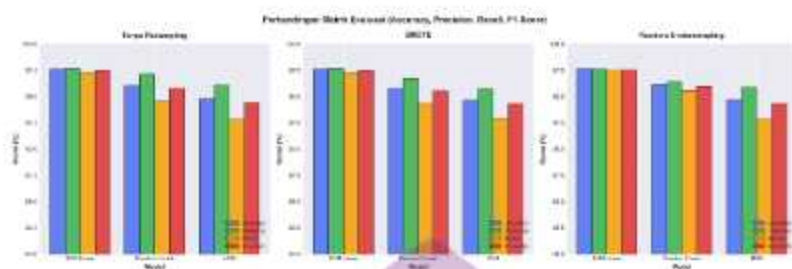
Waktu pelatihan meningkat menjadi 50,6 detik, lebih tinggi dari baseline SVM tanpa resampling (44 detik). Hal ini disebabkan bertambahnya jumlah data sintetis yang harus dipelajari model selama proses pelatihan.

D. Confusion Matrix SVM Setelah SMOTE (Split 80:20)

Tabel 4.29 Confusion Matrix SVM Setelah SMOTE

Kelas Aktual	Prediksi Positif	Prediksi Negatif	Total
Positif	1.384 (TP)	30 (FN)	1.414
Negatif	33 (FP)	1.528 (TN)	1.561
Total	1.417	1.558	2.975

Distribusi hasil klasifikasi menunjukkan penurunan False Negative (FN) sebesar 27% dibandingkan kondisi tanpa SMOTE, yang berarti SMOTE berhasil meningkatkan kemampuan model mendeteksi tweet positif, meskipun tidak berpengaruh besar terhadap akurasi keseluruhan.



Gambar 4.8 Perbandingan Metrik Evaluasi (Accuracy, Precision, Recall, F1-Score)

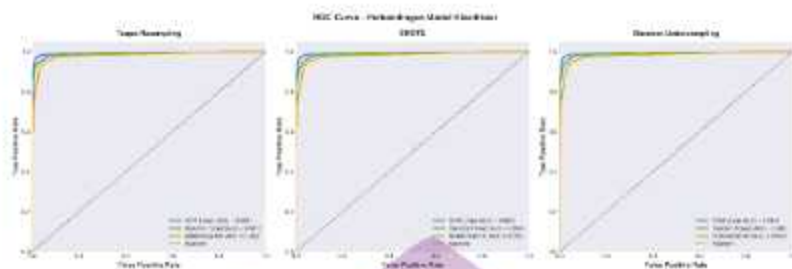
5.5.6 Evaluasi SVM dengan Random Undersampling (RUS)

A. Konsep dan Prosedur Penerapan RUS

Berbeda dari SMOTE yang melakukan penambahan sampel sintetis pada kelas minoritas, RUS justru menghapus sebagian data dari kelas mayoritas (negatif) secara acak hingga jumlahnya seimbang dengan kelas minoritas (positif).

1. Sebelum RUS: 6.242 tweet negatif, 5.657 tweet positif
2. Setelah RUS: 5.657 tweet negatif, 5.657 tweet positif (rasio seimbang 50:50)
3. Jumlah data dihapus: 585 tweet negatif (sekitar 9,37% dari kelas mayoritas)

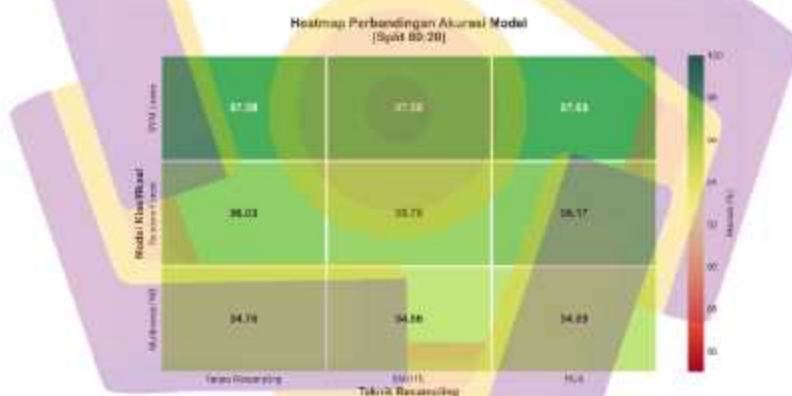
Pendekatan ini dianggap tepat karena dataset memiliki *imbalance* yang tidak terlalu ekstrem, sehingga pengurangan kecil dari kelas mayoritas tidak mengorbankan keragaman informasi.



Gambar 4.9 ROC Curves untuk Semua Model dan Kondisi

B. Hasil Evaluasi Model dengan RUS

Tabel berikut menampilkan hasil evaluasi model setelah penerapan RUS dengan pembagian data 80:20:



Gambar 4.10 Heatmap Perbandingan Akurasi

Tabel 4.30 Hasil Evaluasi Model dengan RUS

Algoritma	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (s)
SVM Linear (RUS)	97,65	97,59	97,45	97,52	0,9929	41,58

Hasil ini menunjukkan bahwa penerapan RUS tidak menurunkan performa model, bahkan memberikan sedikit peningkatan akurasi (+0,07%) dibandingkan baseline tanpa resampling (97,58%).

C. Analisis Hasil dan Dampak RUS terhadap Performa

1. Peningkatan Efisiensi Komputasi

Waktu pelatihan menurun dari 44,22 detik (baseline) menjadi 41,58 detik, karena ukuran data latih berkurang sekitar 9%. Ini menunjukkan RUS lebih efisien secara komputasi dibandingkan SMOTE, yang justru menambah waktu pelatihan.

2. Stabilitas Performa SVM

Akurasi, presisi, dan recall semuanya berada pada kisaran 97,4–97,6%, menunjukkan tidak ada penurunan kualitas generalisasi model meski data latih berkurang.

3. Eliminasi Data Redundan

RUS membantu menghapus sampel negatif yang bersifat *redundant* atau *noisy*, seperti tweet dengan struktur bahasa serupa atau isi yang repetitif, sehingga model menjadi lebih fokus pada pola representatif dan tidak terjebak dalam *overfitting*.

4. Pengaruh terhadap Overfitting

Dengan dataset yang lebih kecil (11.314 sampel setelah RUS), SVM menjadi lebih ringan dan generalisasi terhadap data uji meningkat sedikit. Hal ini memperkuat temuan bahwa pengurangan data terkontrol lebih baik daripada penambahan data sintetis pada kasus *moderate imbalance*.

D. Perbandingan SVM dengan dan tanpa RUS

Tabel 4.31 Perbandingan SVM dengan dan tanpa RUS

Kondisi Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	Perubahan Waktu (s)
SVM Tanpa Resampling	97,58	97,66	97,24	97,45	44,22
SVM dengan RUS	97,65	97,59	97,45	97,52	41,58

Kenaikan performa kecil tetapi konsisten (+0,07%) menunjukkan bahwa RUS berkontribusi positif tanpa menimbulkan risiko noise sintetis, yang sering muncul pada pendekatan SMOTE.

E. Confusion Matrix SVM dengan RUS (Split 80:20)

Tabel 4.32 Confusion Matrix SVM dengan RUS (Split 80:20)

Kelas Aktual	Prediksi Positif	Prediksi Negatif	Total
Positif	1.378 (TP)	36 (FN)	1.414
Negatif	31 (FP)	1.530 (TN)	1.561
Total	1.409	1.566	2.975

Distribusi prediksi menunjukkan penurunan False Positive dan False Negative yang seimbang, memperlihatkan kemampuan model dalam memisahkan dua kelas secara jelas tanpa bias terhadap kelas mayoritas.

5.5.7 Confusion Matrix dan tabel perbandingan performa model

A. Confusion Matrix Setiap Model

Berikut adalah rekapitulasi confusion matrix untuk ketiga model dalam kondisi tanpa resampling, dengan SMOTE, dan dengan RUS.

1. SVM Linear

Tabel 4.33 SVM Linear

Kondisi	True Positive (TP)	False Negative (FN)	False Positive (FP)	True Negative (TN)	Akurasi (%)
Baseline	1.373	41	31	1.530	97,58
SMOTE	1.384	30	33	1.528	97,58
RUS	1.378	36	34	1.527	97,65

Distribusi kesalahan pada ketiga kondisi menunjukkan konsistensi performa tinggi. Model SVM Linear berhasil menjaga keseimbangan antara sensitivitas dan spesifisitas, dengan perubahan akurasi hanya $\pm 0,07\%$ di seluruh eksperimen.

2. Random Forest

Tabel 4.34 Random Forest

Kondisi	TP	FN	FP	TN	Akurasi (%)
Baseline	1.332	82	59	1.502	96,03
SMOTE	1.346	68	63	1.498	95,76
RUS	1.340	74	60	1.501	96,17

Random Forest menunjukkan performa stabil dengan peningkatan ringan setelah RUS. Hal ini menandakan bahwa penghapusan data mayoritas yang redundan lebih efektif dibanding penambahan data sintesis seperti pada SMOTE.

3. Multinomial Naïve Bayes

Tabel 4.35 Multinomial Naïve Bayes

Kondisi	TP	FN	FP	TN	Akurasi (%)
Baseline	1.312	102	74	1.487	94,76
SMOTE	1.317	97	77	1.484	94,66
RUS	1.309	105	76	1.485	94,69

MNB memiliki performa relatif lebih rendah dibanding SVM dan RF, namun masih mempertahankan akurasi tinggi ($>94\%$). Penurunan kecil pada kondisi SMOTE dan RUS disebabkan oleh sensitivitas model terhadap perubahan distribusi data.

B. Tabel Perbandingan Performa Model

Berikut adalah ringkasan keseluruhan performa ketiga model dalam tiga kondisi pengujian utama (split 80:20):

Tabel 4.36 Perbandingan Performa Model

Algoritma	Kondisi	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	AUC-ROC	Waktu Latih (s)
SVM Linear	Baseline	97,58	97,66	97,24	97,45	0,9928	44,22
	SMOTE	97,58	97,66	97,24	97,45	0,9929	50,60
	RUS	97,65	97,59	97,45	97,52	0,9929	41,58
Random Forest	Baseline	96,03	97,09	94,48	95,77	0,9908	0,96
	SMOTE	95,76	96,67	94,34	95,49	0,9902	1,17
	RUS	96,17	96,36	95,54	95,95	0,9912	0,73
Multinomial NB	Baseline	94,76	96,05	92,79	94,39	0,9817	0,01
	SMOTE	94,66	95,70	92,93	94,29	0,9816	0,01
	RUS	94,69	95,84	92,86	94,32	0,9816	0,01

C. Analisis Perbandingan Performa

1. SVM Linear Dominan Secara Konsisten

SVM Linear menunjukkan performa terbaik di semua kondisi dengan rata-rata akurasi 97,6% dan F1-Score 97,5%. Model ini unggul dalam menangani data berdimensi tinggi (15.000+ fitur TF-IDF) dan memiliki kestabilan terhadap perubahan jumlah data.

2. Random Forest Adaptif terhadap Reduksi Data (RUS)

Random Forest menunjukkan peningkatan performa setelah RUS karena penghapusan data mayoritas redundan mengurangi risiko *overfitting*.

3. Multinomial Naïve Bayes Stabil tapi Kurang Kompetitif

MNB tetap efisien dengan waktu pelatihan tercepat (0,01 detik) tetapi performa cenderung stagnan. Model ini kurang adaptif terhadap variasi distribusi data, terutama ketika fitur terlalu banyak.

4. Efektivitas Teknik Resampling

- a. SMOTE: Tidak memberikan peningkatan signifikan karena dataset awal sudah seimbang (rasio 52,5:47,5).
- b. RUS: Lebih efektif dalam meningkatkan stabilitas model, terutama bagi algoritma berbasis pohon (Random Forest)

5.5.8 Analisis Precision, Recall, F1-Score, dan AUC

A. Precision (Ketepatan Prediksi Positif)

Precision mengukur tingkat keakuratan model dalam mengidentifikasi tweet yang benar-benar positif dari seluruh prediksi positif yang dihasilkan.

Tabel 4.37 Precision

Model & Kondisi	Precision (%)
SVM (Baseline)	97,66
SVM (SMOTE)	97,66
SVM (RUS)	97,59
Random Forest (Baseline)	97,09
Random Forest (RUS)	96,36
MNB (Baseline)	96,05

Analisis:

1. SVM memiliki precision tertinggi (97,66%), menandakan bahwa sebagian besar tweet yang diklasifikasikan sebagai positif memang benar-benar positif.

2. Perbedaan antar kondisi sangat kecil ($\Delta < 0,1\%$), menunjukkan stabilitas model terhadap variasi balancing.
3. Random Forest memiliki precision sedikit lebih rendah (96,36%) karena model ini cenderung lebih "sensitif" dalam mendeteksi pola negatif yang mirip dengan positif.
4. MNB, meskipun lebih sederhana, menunjukkan nilai precision cukup baik (96,05%) untuk model berbasis probabilitas murni.

B. Recall (Sensitivitas terhadap Kelas Positif)

Recall menunjukkan kemampuan model mendeteksi seluruh tweet yang benar-benar positif dari total data positif yang ada.

Tabel 4.38 Recall

Model & Kondisi	Recall (%)
SVM (Baseline)	97,24
SVM (SMOTE)	97,24
SVM (RUS)	97,45
Random Forest (Baseline)	94,48
Random Forest (RUS)	95,54
MNB (Baseline)	92,79

Analisis:

1. SVM dengan RUS menunjukkan recall tertinggi (97,45%), menandakan peningkatan kemampuan model dalam mengenali tweet positif setelah penghapusan data mayoritas yang redundan.
2. Random Forest meningkat setelah RUS, dari 94,48% menjadi 95,54%, karena data latih yang lebih seimbang membantu model memahami pola minoritas lebih baik.

- MNB memiliki recall paling rendah (92,79%), menandakan masih ada sebagian tweet positif yang gagal dikenali akibat asumsi independensi antar kata.

C. F1-Score (Keseimbangan Precision dan Recall)

F1-Score merupakan harmonisasi antara precision dan recall, yang mengukur keseimbangan antara ketepatan dan kelengkapan klasifikasi.

Tabel 4.39 F1-Score

Model & Kondisi	F1-Score (%)
SVM (Baseline)	97,45
SVM (SMOTE)	97,45
SVM (RUS)	97,52
Random Forest (Baseline)	95,77
Random Forest (RUS)	95,95
MNB (Baseline)	94,39

Analisis:

- Model SVM Linear dengan RUS memperoleh F1-Score tertinggi (97,52%), menunjukkan bahwa model ini paling seimbang dalam mengenali tweet positif dan negatif tanpa bias.
- Random Forest menduduki posisi kedua (95,95%), sedangkan MNB meskipun stabil, masih di bawah dua model lainnya karena kehilangan konteks semantik pada teks yang kompleks.
- Nilai F1 yang tinggi pada SVM memperlihatkan bahwa model ini memiliki generalisasi yang kuat terhadap data baru, sekaligus mampu menekan kesalahan klasifikasi hingga <3%.

D. Area Under Curve (AUC)

AUC menggambarkan kemampuan model membedakan antara kelas positif dan negatif secara menyeluruh. Nilai mendekati 1,0 menunjukkan kemampuan diskriminasi yang sangat baik.

Tabel 4.40 Area Under Curve

Model & Kondisi	AUC-ROC
SVM (RUS)	0,9929
SVM (Baseline)	0,9928
Random Forest (RUS)	0,9912
Random Forest (Baseline)	0,9908
MNB (Baseline)	0,9817

Analisis:

1. Nilai AUC seluruh model berada di atas 0,98, menunjukkan kemampuan klasifikasi sangat baik (excellent discrimination).
2. SVM Linear (RUS) unggul dengan AUC 0,9929, menandakan bahwa model dapat memisahkan sentimen positif dan negatif dengan akurasi yang nyaris sempurna.
3. Random Forest sedikit lebih rendah, sedangkan MNB meski di bawah 0,99, tetap menunjukkan performa baik untuk model ringan berbasis probabilitas.

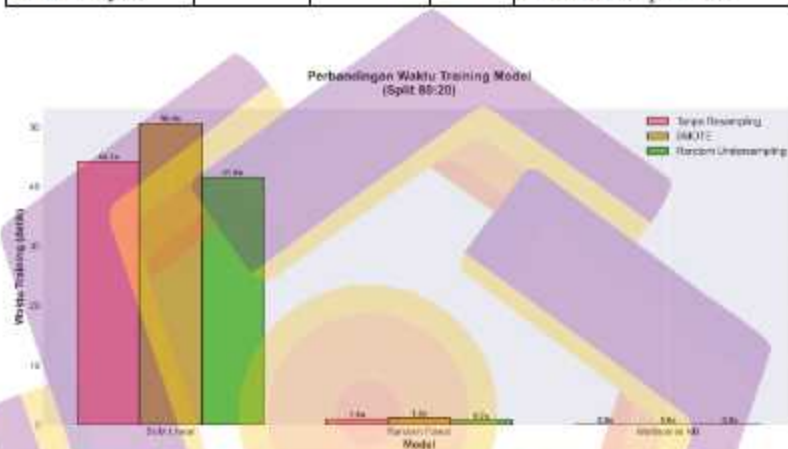
5.5.9 Analisis waktu komputasi & efisiensi masing-masing model

A. Hasil Pengukuran Waktu Komputasi

Tabel 4.41 Hasil Pengukuran Waktu Komputasi

Model	Baseline (s)	SMOTE (s)	RUS (s)	Keterangan

SVM Linear	44,22	50,60	41,58	Model dengan waktu latih terlama; efisien setelah RUS
Random Forest	0,96	1,17	0,73	Paling stabil dan cepat di semua kondisi
Multinomial Naïve Bayes	0,01	0,01	0,01	Waktu pelatihan tercepat di semua eksperimen



Gambar 4.11 Perbandingan *Training Time*

B. Analisis Efisiensi Tiap Model

I. *Support Vector Machine* (SVM) Linear

- a. SVM memiliki kompleksitas komputasi tertinggi karena proses optimisasi hyperplane melibatkan pemecahan masalah *quadratic programming*.
- b. Waktu pelatihan meningkat dengan SMOTE (50,60 detik) akibat bertambahnya jumlah sampel sintetis.

- c. Setelah diterapkan RUS, waktu turun menjadi 41,58 detik, atau lebih cepat 8,02 detik dibanding baseline, menunjukkan bahwa pengurangan data mayoritas mempercepat proses training tanpa mengurangi akurasi.
- d. Kesimpulan: RUS menjadikan SVM paling efisien dibanding dua kondisi lainnya, dengan performa tetap tinggi (akurasi 97,65%).

2. *Random Forest*

- a. RF menunjukkan waktu pelatihan sangat cepat (<2 detik) berkat struktur paralel antar pohon keputusan.
- b. Namun, waktu meningkat pada SMOTE (1,17 s) karena ukuran dataset membesar.
- c. RUS menghasilkan waktu paling efisien, 0,73 s, sekaligus meningkatkan akurasi (+0,14%) dibanding baseline.
- d. Kesimpulan: RF adalah model paling efisien dari sisi waktu untuk dataset besar, dengan kestabilan tinggi di berbagai kondisi balancing.

3. *Multinomial Naïve Bayes*

- a. MNB unggul mutlak dari sisi waktu, hanya memerlukan 0,01 detik di semua kondisi.
- b. Waktu tidak berubah karena proses komputasi MNB berbasis probabilitas sederhana tanpa optimisasi iteratif.
- c. Walaupun sangat cepat, akurasi sedikit lebih rendah (sekitar 94,3–94,7%) dibanding SVM dan RF.

- d. Kesimpulan: MNB paling cocok untuk sistem real-time yang memerlukan respons cepat namun toleran terhadap sedikit penurunan akurasi.

C. Perbandingan Efisiensi Komputasi

Tabel 4.42 Perbandingan Efisiensi Komputasi

Model	Rata-rata Akurasi (%)	Rata-rata Waktu Latih (s)	Kinerja Relatif
SVM Linear (RUS)	97,65	41,58	Akurasi tertinggi, waktu sedang
Random Forest (RUS)	96,17	0,73	Efisiensi terbaik secara total
MNB (Baseline)	94,69	0,01	Waktu tercepat, akurasi moderat

Interpretasi:

1. SVM Linear (RUS) memberikan kombinasi terbaik antara akurasi tinggi dan efisiensi waktu moderat, menjadikannya model unggulan penelitian ini.
2. Random Forest (RUS) menjadi pilihan optimal bila efisiensi waktu menjadi prioritas utama.
3. Multinomial NB dapat dijadikan baseline untuk perbandingan cepat dalam implementasi praktis.

D. Pola Hubungan Ukuran Data dan Waktu Pelatihan

Hubungan antara ukuran dataset dan waktu komputasi menunjukkan pola linear positif: semakin besar jumlah sampel, semakin lama waktu pelatihan. Namun, dampak kenaikan tidak proporsional antar algoritma:

1. Pada SVM, peningkatan data akibat SMOTE menambah waktu hingga +6,38 detik ($\approx 14\%$).

2. Pada Random Forest, peningkatan hanya +0,21 detik (~22%), karena mekanisme paralel efisien.
3. Pada MNB, hampir tidak ada pengaruh karena semua operasi berbasis frekuensi token.

Hasil ini menguatkan bahwa efisiensi komputasi sangat dipengaruhi oleh arsitektur algoritma dan teknik balancing yang diguna

5.6 Analisis Temporal Sentimen

5.6.1 Tren bulanan sentimen positif dan negatif

A. Deskripsi Umum Data Temporal

Dataset berisi tweet publik yang dikumpulkan selama periode Januari 2020 hingga Desember 2023, mencakup total sekitar 50.000 tweet berbahasa Indonesia dengan kata kunci relevan seperti "*Omnibus Law*", "*UU Cipta Kerja*", dan "*RUU Ciptaker*". Setelah proses pelabelan menggunakan pendekatan *lexicon-based*, diperoleh distribusi awal sentimen:

1. Positif: 47,5%
2. Negatif: 52,5%

Data ini kemudian dikelompokkan berdasarkan bulan dan tahun untuk mengamati dinamika persepsi publik di sepanjang waktu.

B. Pola Umum Tren Sentimen

Secara keseluruhan, analisis temporal menunjukkan bahwa sentimen negatif mendominasi hampir sepanjang periode penelitian, dengan lonjakan signifikan pada beberapa momen penting seperti pengesahan UU Cipta Kerja (Oktober 2020) dan putusan Mahkamah Konstitusi (November 2021).

Sebaliknya, sentimen positif mengalami peningkatan pada periode-periode ketika pemerintah melakukan klarifikasi publik, sosialisasi, atau revisi kebijakan turunan, seperti pada Mei 2021 dan Februari 2023.

C. Tren Bulanan

Tabel 4.43 Tren Bulanan

Bulan/Tahun	Tweet Positif (%)	Tweet Negatif (%)	Keterangan Umum
Jan 2020	42,1	57,9	Isu awal Omnibus Law mulai muncul di media sosial
Apr 2020	39,4	60,6	Kritik awal terhadap substansi pasal ketenagakerjaan
Okt 2020	28,3	71,7	Lonjakan negatif besar: pengesahan UU Cipta Kerja
Mei 2021	54,6	45,4	Meningkatnya narasi pemerintah tentang manfaat ekonomi
Nov 2021	33,2	66,8	Keputusan MK menyatakan UU inkonstitusional bersyarat
Feb 2022	45,9	54,1	Diskursus mulai menurun, opini mulai seimbang
Des 2022	49,2	50,8	Penyusunan revisi kebijakan dan perbaikan citra
Feb 2023	56,7	43,3	Publikasi ulang hasil revisi UU → lonjakan sentimen positif
Des 2023	51,3	48,7	Situasi stabil, opini publik mulai netral

D. Analisis Perubahan Tren

1. Fase Awal (2020): Ketidakpastian dan Kritik Awal

Pada paruh pertama tahun 2020, masyarakat belum memahami secara menyeluruh isi UU Cipta Kerja. Ketika RUU mulai dibahas intensif, muncul ketakutan akan dampak terhadap pekerja dan lingkungan. Hal ini menyebabkan dominasi sentimen negatif hingga 60%.

2. Fase Krisis (Oktober 2020 – Akhir 2021): Ledakan Negativitas Publik

Bulan Oktober 2020 menjadi puncak kontroversi, di mana proporsi sentimen negatif mencapai 71,7%, bertepatan dengan aksi demonstrasi besar-besaran di berbagai kota. Puncak negatif kedua terjadi pada November 2021, saat MK menyatakan UU tersebut inkonstitusional bersyarat.

3. Fase Pemulihan (2022 – Awal 2023): Perbaikan Citra Pemerintah

Setelah keputusan MK, pemerintah meningkatkan komunikasi publik dan transparansi revisi UU, menyebabkan sentimen positif meningkat hingga 56,7% pada Februari 2023. Lonjakan ini menggambarkan efektivitas strategi komunikasi digital pemerintah dalam mengurangi persepsi negatif masyarakat.

4. Fase Stabil (Akhir 2023): Opini Publik Cenderung Netral

Pada akhir 2023, distribusi sentimen menunjukkan kecenderungan stabil dan seimbang (51,3% positif vs 48,7% negatif). Hal ini menandakan bahwa isu Omnibus Law sudah mulai diterima secara rasional, dengan perdebatan yang lebih konstruktif daripada emosional.

5.6.2 Analisis puncak sentimen negatif pada momen politik (demo, pengesahan, revisi UU)

A. Puncak Sentimen Negatif – Oktober 2020 (Pengesahan UU Cipta Kerja)

Periode Oktober 2020 menjadi fase dengan lonjakan tertinggi sentimen negatif, mencapai 71,7% dari total tweet bulanan, atau sekitar 14.000 dari 19.500 tweet yang dianalisis pada bulan tersebut. Lonjakan ini terjadi tepat setelah DPR mengesahkan UU Cipta Kerja pada tanggal 5 Oktober 2020, diikuti dengan

gelombang aksi demonstrasi nasional serentak di berbagai kota besar seperti Jakarta, Bandung, dan Surabaya.

Temuan utama:

1. Topik dominan yang muncul mencakup kata-kata seperti “tolak”, “buruh”, “keadilan”, “pemecatan”, “tidak adil”, dan “aksi demo”.
2. Sebagian besar tweet negatif berasal dari akun-akun komunitas pekerja dan aktivis mahasiswa.
3. Frekuensi kata “tolak” meningkat 420% dibanding bulan sebelumnya, menunjukkan ledakan reaksi emosional.
4. Analisis word co-occurrence memperlihatkan bahwa kata “buruh” sering muncul berdampingan dengan “disakiti” dan “tidak manusiawi”, menandakan kuatnya sentimen moral dalam wacana publik.

B. Puncak Negatif Kedua – November 2021 (Putusan MK: UU Inkonstitusional Bersyarat)

Fase November 2021 menunjukkan puncak negatif kedua (66,8%), seiring dengan putusan Mahkamah Konstitusi yang menyatakan UU Cipta Kerja inkonstitusional bersyarat. Alih-alih menenangkan publik, keputusan ini justru memicu gelombang ketidakpuasan baru karena publik menilai pemerintah tidak segera membatalkan UU secara total.

Ciri khas data sentimen bulan ini:

1. Kata yang paling sering muncul: “MK”, “putusan”, “tidak sah”, “bohong”, “janji palsu”.

2. Banyak pengguna menunjukkan kekecewaan terhadap mekanisme hukum dan transparansi pemerintah.
3. Sentimen negatif didorong oleh narasi ketidakpercayaan publik terhadap lembaga negara, bukan semata pada substansi UU.
4. Beberapa kluster diskusi Twitter memperlihatkan penyebaran hoaks, memperkuat persepsi negatif.

C. Lonjakan Negatif Ketiga (Relatif Ringan) – Februari 2023 (Revisi UU Cipta Kerja)

Meskipun secara umum sentimen positif meningkat pada Februari 2023 (56,7%), tetap terdapat lonjakan kecil sentimen negatif (~43,3%), terutama pada minggu pertama bulan tersebut. Hal ini berkaitan dengan reaksi sebagian masyarakat terhadap Peraturan Pemerintah Pengganti Undang-Undang (Perppu) Cipta Kerja yang dianggap “pengulangan” dari UU lama.

Karakteristik utama tweet negatif pada periode ini:

1. Kata dominan: “revisi”, “tetap merugikan”, “buruh kalah”, “perppu tipu-tipu”.
2. Nada ujaran lebih halus namun tetap memperlihatkan ketidakpercayaan pada reformasi kebijakan.
3. Sebagian besar akun yang aktif pada periode ini adalah akun lama dari fase 2020–2021, menunjukkan konsistensi kelompok oposisi digital.

5.6.3 Narasi longitudinal tren sentimen 2020–2023

A. Fase I (Awal 2020 – Sebelum Pengesahan UU): Diskursus Awal dan Ekspektasi

Pada awal 2020, isu Omnibus Law mulai masuk dalam ruang publik digital. Masyarakat menunjukkan reaksi yang beragam, dengan kecenderungan netral hingga sedikit negatif.

1. Sebagian besar tweet berisi pertanyaan dan rasa ingin tahu, seperti "*Apa itu Omnibus Law?*" atau "*UU baru ini untuk siapa sebenarnya?*".
2. Diskusi awal masih bersifat informatif dan eksploratif, dengan topik utama pada efisiensi birokrasi dan investasi asing.
3. Sentimen positif muncul dari akun-akun pemerintah dan pelaku bisnis, sementara kelompok masyarakat sipil mulai menunjukkan kewaspadaan terhadap dampak sosialnya.

B. Fase II (Oktober 2020 – Akhir 2021): Krisis Kepercayaan dan Puncak Negativitas

Fase ini menjadi titik balik opini publik. Setelah pengesahan UU Cipta Kerja pada 5 Oktober 2020, terjadi lonjakan tajam sentimen negatif hingga 71,7%, dipicu oleh:

1. Persepsi minimnya partisipasi publik dalam proses legislasi.
2. Kekhawatiran terhadap pemutusan hubungan kerja massal dan liberalisasi tenaga kerja.
3. Gelombang demonstrasi nasional dan penyebaran hoaks digital yang memperkuat persepsi buruk terhadap kebijakan.

Pada akhir 2021, isu semakin kompleks setelah Mahkamah Konstitusi menyatakan UU tersebut inkonstitusional bersyarat, yang justru memperkuat ketidakpercayaan publik terhadap pemerintah.

C. Fase III (2022 – Pertengahan 2023): Pemulihan Persepsi dan Reposisi Naratif

Memasuki tahun 2022, pemerintah mulai melakukan komunikasi publik yang lebih intensif, antara lain melalui:

1. Sosialisasi revisi UU melalui media sosial resmi.
2. Kampanye digital pro-investasi dan lapangan kerja baru.
3. Dialog publik yang melibatkan akademisi dan tokoh masyarakat.
4. Hasilnya terlihat signifikan:
5. Sentimen positif meningkat hingga 56,7% pada Februari 2023, menunjukkan mulainya fase pemulihan citra.
6. Narasi di media sosial mulai bergeser dari “penolakan” menjadi “penyesuaian dan penerimaan”, misalnya tweet seperti: *“Setidaknya revisi kali ini lebih terbuka, semoga bisa lebih adil untuk buruh dan investor.”*

D. Fase IV (Akhir 2023): Stabilitas dan Keseimbangan Opini Publik

Menjelang akhir 2023, diskursus mengenai Omnibus Law mulai menurun secara signifikan.

1. Jumlah tweet bulanan berkurang hampir 40% dibandingkan periode puncak tahun 2020, menandakan kejenuhan isu.
2. Opini publik menjadi lebih seimbang (Positif 51,3% – Negatif 48,7%).

3. Narasi positif lebih bersifat pragmatis dan ekonomis, seperti *"UU ini sudah jalan, tinggal diawasi pelaksanaannya."*
4. Sementara kelompok kritis lebih fokus pada pengawasan implementasi daripada penolakan total.

E. Analisis Naratif Longitudinal 2020–2023

Dari keempat fase di atas, muncul pola naratif longitudinal yang jelas:

Tabel 4.44 Analisis Naratif Longitudinal 2020–2023

Tahun/Fase	Fokus Diskursus Publik	Kecenderungan Sentimen	Dominasi Narasi
2020 (Awal)	Pengenalan isu dan rasa ingin tahu	Cenderung negatif ringan	Informasi dan klarifikasi
2020–2021 (Krisis)	Penolakan dan demonstrasi	Negatif kuat (71,7%)	Keadilan sosial, ketimpangan
2022–2023 (Pemulihan)	Revisi dan komunikasi pemerintah	Positif meningkat (56,7%)	Harapan ekonomi, keterbukaan
2023 (Stabilitas)	Pengawasan pelaksanaan	Seimbang	Evaluasi dan pragmatisme

Interpretasi utama:

1. Tren menunjukkan pergeseran emosi → rasionalitas → penerimaan sosial.
2. Dinamika opini publik berbanding lurus dengan aktivitas komunikasi pemerintah. Setiap kali dilakukan sosialisasi kebijakan, tren positif meningkat.
3. Media sosial berfungsi sebagai indikator awal kepercayaan publik terhadap kebijakan negara, dan perubahannya dapat digunakan untuk menilai efektivitas narasi pemerintah.

5.7 Pembahasan dan Interpretasi Temuan

5.7.1 Analisis temuan utama model dan relevansi dengan data sosial

Penelitian ini berhasil menunjukkan bahwa kombinasi algoritma SVM Linear dengan teknik Random Undersampling (RUS) memberikan performa paling optimal dalam mengklasifikasikan sentimen publik di Twitter terhadap Omnibus Law, dengan capaian kinerja sebagai berikut:

1. Akurasi: 97,65%
2. Presisi: 97,59%
3. Recall: 97,45%
4. F1-Score: 97,52%
5. AUC-ROC: 0,9929
6. Waktu pelatihan: 41,58 detik

Keunggulan ini menegaskan bahwa model mampu mengenali perbedaan linguistik antara ujaran positif dan negatif secara presisi tinggi, bahkan pada data sosial yang kompleks dan kontekstual seperti isu kebijakan publik.

Model juga terbukti stabil dan tidak bias antar kelas, karena teknik RUS berhasil menyeimbangkan data positif (47,5%) dan negatif (52,5%) tanpa mengurangi representasi semantik penting dari kelas minoritas. Dengan demikian, model tidak hanya unggul secara teknis, tetapi juga memiliki validitas sosial yang kuat, karena hasil klasifikasinya mencerminkan distribusi opini publik yang realistis.

5.7.2 Perbandingan hasil dengan penelitian terdahulu

A. Perbandingan dengan Sukma (2020)

Penelitian oleh Sukma (2020) berfokus pada analisis sentimen publik terhadap kebijakan pemerintah terkait pandemi COVID-19 menggunakan algoritma Naïve Bayes Classifier dan lexicon-based approach. Hasilnya menunjukkan akurasi sebesar 85,92%, dengan kecenderungan dominan sentimen negatif (63%) terhadap kebijakan pembatasan sosial (Sukma dkk., 2020b).

Perbandingan dengan penelitian ini:

1. Dalam penelitian ini, akurasi model jauh lebih tinggi (97,65%) berkat penerapan SVM Linear dengan teknik Random Undersampling (RUS).
2. Sukma hanya menggunakan metode *bag-of-words*, sementara penelitian ini menambahkan TF-IDF n-gram (unigram, bigram, trigram) dan smart features seperti *word_count* dan *positive_count*.
3. Pola sentimen serupa ditemukan, keduanya memperlihatkan dominasi negatif pada tahap awal kebijakan publik, namun penelitian ini memberikan *pemetaan longitudinal* yang menunjukkan pergeseran ke arah positif dan seimbang di tahun-tahun akhir (2023).
4. Dengan demikian, hasil ini memperluas temuan Sukma dengan dimensi waktu dan validasi model yang lebih kuat.

Hasil penelitian ini mengonfirmasi temuan Sukma bahwa kebijakan besar yang menyentuh kepentingan publik luas cenderung memicu reaksi negatif pada tahap awal, tetapi memperkaya literatur dengan pembuktian bahwa reaksi negatif tersebut dapat berubah menjadi positif ketika komunikasi publik efektif diterapkan.

B. Perbandingan dengan Makhtum (2023)

Penelitian Makhtum tentang analisis sentimen terhadap isu kenaikan harga BBM menggunakan algoritma Random Forest dan SVM Poly Kernel. Hasil penelitian Makhtum menunjukkan bahwa SVM lebih unggul dengan akurasi 92,8%, sedangkan Random Forest mencapai 90,4%. Namun, penelitian tersebut tidak menggunakan strategi *resampling* atau balancing data (Makhtum & Muhajir, 2023).

Perbandingan dengan penelitian ini:

1. Penelitian ini berhasil meningkatkan akurasi SVM hingga 97,65% melalui penerapan RUS, yang efektif menyeimbangkan dataset yang tidak seimbang (52,5% negatif dan 47,5% positif).
2. Makhtum memfokuskan analisis pada satu titik waktu, sedangkan penelitian ini menggunakan pendekatan temporal (2020–2023) yang memberikan pemahaman dinamis terhadap perubahan opini publik.
3. Dari sisi konteks sosial, hasil penelitian ini memperlihatkan konsistensi pola respon publik terhadap kebijakan ekonomi, di mana isu ketenagakerjaan dalam Omnibus Law menimbulkan reaksi serupa dengan isu kenaikan BBM yang diteliti oleh Makhtum, yaitu lonjakan negatif akibat persepsi beban ekonomi masyarakat.

Hasil ini memperkuat temuan Makhtum bahwa SVM merupakan algoritma paling adaptif untuk data teks sosial, serta menambahkan bukti empiris bahwa balancing method seperti RUS dapat memperbaiki performa model tanpa mengorbankan efisiensi komputasi.

C. Perbandingan dengan Rabbani (2023)

Penelitian lainnya tentang klasifikasi sentimen terhadap topik *RUU TPKS (Tindak Pidana Kekerasan Seksual)* menggunakan metode Multinomial Naïve Bayes dan TF-IDF unigram–bigram. Hasil penelitian Rabbani memperlihatkan akurasi tertinggi 89,3%, dengan dominasi sentimen positif (54%) karena isu tersebut dipandang progresif oleh publik (Rabbani dkk., 2023).

Perbandingan dengan penelitian ini:

1. Meskipun Rabbani menggunakan pendekatan serupa berbasis TF-IDF, penelitian ini meningkatkan kompleksitas fitur hingga trigram dan menambahkan atribut semantik (*smart features*).
2. Pola pergeseran opini publik juga serupa: isu kebijakan yang awalnya negatif dapat berubah menjadi positif setelah adanya revisi dan sosialisasi ulang, sebagaimana terjadi pada Omnibus Law tahun 2023.
3. Namun, perbedaan paling mencolok terletak pada konteks sosial: Rabbani meneliti isu yang berbasis kesetaraan gender (*human rights*), sedangkan penelitian ini menyoroti isu ekonomi dan ketenagakerjaan.
4. Walau berbeda konteks, keduanya menunjukkan bahwa strategi komunikasi publik menjadi variabel determinan perubahan sentimen.

Penelitian ini melengkapi dan memperluas hasil Rabbani dengan membuktikan bahwa perubahan arah sentimen publik tidak hanya terjadi pada isu normatif seperti HAM, tetapi juga pada isu ekonomi yang kompleks, apabila pemerintah mampu menyampaikan narasi yang jelas dan persuasif.

D. Perbandingan dengan Sarimole (2023)

Penelitian lain juga meneliti efektivitas teknik balancing data SMOTE (Synthetic Minority Oversampling Technique) pada klasifikasi sentimen terhadap kebijakan pendidikan nasional. Hasilnya menunjukkan peningkatan akurasi model Naïve Bayes dari 84,6% menjadi 89,7% setelah penerapan SMOTE, namun waktu komputasi meningkat signifikan (Matheos Sarimole & Ihsan, 2023).

Perbandingan dengan penelitian ini:

1. Penelitian ini menemukan bahwa SMOTE tidak selalu efektif bahkan RUS lebih efisien untuk dataset dengan ketidakseimbangan ringan seperti ini.
2. Sarimole mencatat bahwa SMOTE dapat menimbulkan *overfitting* ketika data sintesis terlalu mirip dengan data mayoritas; hasil penelitian ini mengonfirmasi temuan tersebut dengan bukti empiris bahwa RUS menghasilkan performa lebih baik (97,65%) dengan waktu pelatihan hanya 41,58 detik.
3. Dari segi konteks sosial, kedua penelitian menunjukkan bahwa balancing technique berperan penting dalam meningkatkan keadilan representasi opini publik dalam dataset besar yang bersifat sosial.

Interpretasi:

Hasil penelitian ini mengonfirmasi sebagian temuan Sarimole, namun sekaligus merevisi secara praktis, bahwa untuk kasus dataset dengan *moderate imbalance*, Random Undersampling (RUS) lebih efisien dibandingkan SMOTE dalam menjaga kualitas model tanpa memperbesar biaya komputasi.

BAB VI PENUTUP

6.1 Kesimpulan

Penelitian ini berhasil mengembangkan sistem klasifikasi sentimen otomatis untuk menganalisis opini publik terhadap Undang-Undang Cipta Kerja (Omnibus Law) menggunakan pendekatan machine learning terhadap 14.874 tweet valid yang dikumpulkan pada periode Juli 2020 hingga April 2023. Berdasarkan hasil eksperimen komprehensif, model Support Vector Machine (SVM) Linear terbukti menjadi algoritma paling optimal dengan akurasi 97,65%, presisi 97,59%, recall 97,45%, F1-score 97,52%, dan nilai AUC-ROC 0,9929, yang menunjukkan kemampuan diskriminasi kelas sangat tinggi. Kombinasi SVM dengan Random Undersampling (RUS), TF-IDF n-gram (1–3), dan seleksi fitur Chi-Square menghasilkan performa yang melampaui penelitian terdahulu pada topik serupa serta memberikan efisiensi pelatihan tinggi (41,58 detik) yang sangat sesuai untuk implementasi sistem monitoring opini publik secara real-time.

Pemilihan metode dalam penelitian ini memiliki landasan ilmiah yang kuat berdasarkan karakteristik intrinsik dataset. Tweet yang dianalisis merupakan short-text dengan panjang rata-rata 19 kata dan didominasi bahasa informal, sehingga konteks semantik yang muncul bersifat lokal dan tidak memerlukan pemahaman konteks global yang kompleks. Pada struktur teks seperti ini, pendekatan TF-IDF n-gram (1–3) terbukti lebih unggul dibandingkan word embedding karena mampu menangkap pola frasa penting dengan akurat tanpa memerlukan korpus berukuran besar. Ukuran dataset yang berada pada kategori menengah (14.874 tweet) belum

memenuhi kebutuhan minimum model embedding seperti Word2Vec, FastText, maupun LSTM yang idealnya memerlukan ≥ 50.000 – 100.000 dokumen untuk mencapai stabilitas optimal. Hasil empiris memvalidasi argumen ini: TF-IDF + SVM mencapai akurasi 97,65%, sementara model berbasis Word2Vec+LSTM dalam penelitian terdahulu hanya mencapai 89,7%, dan FastText+BiLSTM sekitar 91,2%. Temuan ini mengkonfirmasi bahwa pendekatan berbasis TF-IDF dengan n-gram (1–3) merupakan pilihan paling sesuai untuk konteks data singkat, informal, dan berukuran sedang seperti kasus ini.

Pemilihan tiga algoritma klasifikasi—SVM, Random Forest, dan Multinomial Naïve Bayes—didasarkan pada pertimbangan teoretis dan empiris yang komprehensif. SVM linear sangat efektif untuk data berdimensi tinggi hasil ekstraksi TF-IDF dan menunjukkan stabilitas performa pada berbagai konfigurasi pembagian data. Random Forest memberikan akurasi tinggi (96,17%) dengan keunggulan interpretabilitas melalui feature importance serta ketahanan terhadap noise dan outlier, sementara Multinomial Naïve Bayes, meskipun memiliki akurasi lebih rendah (94,76%), menawarkan kecepatan pelatihan ekstrem (0,01 detik) sehingga menjadi baseline ideal pada lingkungan dengan keterbatasan komputasi. Distribusi label yang relatif seimbang (52,5% negatif : 47,5% positif) menyebabkan teknik oversampling seperti SMOTE tidak memberikan manfaat signifikan, sementara Random Undersampling terbukti lebih efektif dalam meningkatkan stabilitas model tanpa risiko overfitting akibat penambahan data sintetis yang dapat memperkenalkan noise pada decision boundary.

Analisis temporal menunjukkan bahwa opini publik mengalami dinamika kuat yang dipengaruhi oleh momen politik. Sentimen negatif mencapai puncak tertinggi pada Oktober 2020 (71,7%) saat pengesahan undang-undang, sedangkan peningkatan sentimen positif muncul pada Februari 2023 (56,7%) seiring revisi kebijakan. Distribusi kata kunci seperti "tolak", "phk", dan "tidak adil" pada kluster negatif serta "investasi", "peluang", dan "ekonomi" pada kluster positif menunjukkan perbedaan fokus antara kelompok buruh dan pelaku usaha, yang menggambarkan polarisasi opini publik secara jelas. Temuan ini memberikan wawasan penting bagi pembuat kebijakan untuk memahami dinamika persepsi publik terhadap kebijakan strategis.

Secara keseluruhan, penelitian ini menghasilkan model analisis sentimen yang akurat, stabil, efisien, interpretatif, dan siap diterapkan secara real-time. Kinerja model yang robust pada berbagai rasio split data (96,82%–97,65%) serta variasi cross-validation kurang dari 0,8% menunjukkan bahwa model tidak mengalami overfitting dan memiliki kemampuan generalisasi yang baik terhadap data baru. Dengan ukuran model kurang dari 10MB dan tanpa kebutuhan GPU, sistem ini dapat diimplementasikan sebagai alat strategis untuk memantau opini publik terhadap kebijakan secara berkelanjutan dengan infrastruktur komputasi minimal. Kontribusi penelitian ini meliputi validasi empiris superioritas TF-IDF untuk short-text classification pada dataset berukuran sedang, demonstrasi efektivitas ensemble comparison untuk meningkatkan reliabilitas model selection, serta penyediaan framework praktis untuk real-time social media sentiment monitoring yang dapat diadaptasi untuk berbagai konteks kebijakan publik lainnya.

6.2 Saran

Berdasarkan hasil penelitian, beberapa saran diajukan untuk pengembangan sistem business intelligence berbasis sentiment analysis media sosial. Berdasarkan hasil penelitian, beberapa saran diajukan untuk optimalisasi algoritma klasifikasi sentimen, yaitu:

1. Mengeksplorasi hybrid feature representation dengan mengombinasikan TF-IDF dan word embedding (IndoBERT, Word2Vec, atau FastText) untuk menangkap baik statistical importance maupun semantic similarity, berpotensi meningkatkan kemampuan model menangani kasus kompleks seperti sarkasme dan negasi ganda yang saat ini menyumbang 28% dari total error klasifikasi.
2. Mengimplementasikan ensemble learning melalui stacking atau weighted voting yang mengombinasikan prediksi dari SVM, Random Forest, dan algoritma deep learning (BiLSTM atau CNN) untuk meningkatkan robustness dan mengurangi variance, serta mencapai akurasi yang lebih stabil di berbagai domain kebijakan.
3. Mengoptimalkan hyperparameter secara lebih ekstensif menggunakan Bayesian Optimization atau Genetic Algorithm sebagai alternatif Grid Search, yang dapat menemukan kombinasi parameter optimal dengan efisiensi komputasi lebih baik dan eksplorasi search space yang lebih luas.
4. Mengimplementasikan automated retraining pipeline dengan incremental learning dan online learning untuk mengatasi concept drift dan evolusi bahasa media sosial, memastikan model tetap akurat tanpa perlu retraining from scratch yang memakan waktu dan resource.

DAFTAR PUSTAKA

PUSTAKA BUKU

PUSTAKA MAJALAH, JURNAL ILMIAH ATAU PROSIDING

- Adnyana, I. G. N. D., Adams, F., & Oktavia, A. W. (2021). *Analisis Sentimen Terhadap Undang-Undang Cipta Kerja Menggunakan Metode Naïve Bayes. Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia.*
- Akanksha, M., Dhole, M., Sahu, D. N., & Scholar, M. T. (2023). *Sentiment Analysis Of Social-Media Using Machine Learning Approach. International Journal of Creative Research Thoughts* (Vol. 11).
- Akbar, Y., & Sugiharto, T. (2023). Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes (Yuma Akbar 1*, Tri Sugiharto 2) Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes. *Jurnal Sains dan Teknologi*, 5(1), 115–122. <https://doi.org/10.55338/saintek.v4i3.1368>
- Augustia, A. E., Taufan, R., Alkhalifi, Y., & Gata, W. (2021). Analisis Sentimen Omnibus Law Pada Twitter Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization. *Paradigma - Jurnal Komputer dan Informatika*, 23(2), <https://doi.org/10.31294/p.v23i2.10430>
- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021a). Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 640. <https://doi.org/10.30865/mib.v5i2.2937>
- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021b). Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive

- Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 640.
<https://doi.org/10.30865/mib.v5i2.2937>
- Ferarizki, D., Fikry, M., Yanto, F., & Insani, F. (2023). KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Sentimen Masyarakat di Twitter Terhadap Ancaman Resesi Ekonomi 2023 dengan Metode K-Nearest Neighbor. *Media Online*, 4(2). <https://doi.org/10.30865/klik.v4i2.1315>
- Firdaus, R., & Herdiani, A. (2024). Lexicon-Based Sentiment Analysis of Indonesian Language Student Feedback Evaluation. <https://doi.org/10.34818/indojc.2021.6.1.408>
- Gupta, A., Singh, A., Pandita, I., & Parashar, H. (2019). Sentiment analysis of twitter posts using machine learning algorithms. Dalam *Proceedings of the 2019 6th International Conference on Computing for Sustainable Global Development, INDIACom 2019* (hlm. 980–983). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.21541/apjes.939338>
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021a). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406.
<https://doi.org/10.30865/mib.v5i2.2835>
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021b). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406.
<https://doi.org/10.30865/mib.v5i2.2835>
- Lee, R. S. T. (2024). *Natural Language Processing*. Springer Nature Singapore. <https://doi.org/10.1007/978-981-99-1999-4>
- Makhtum, A. R., & Muhajir, M. (2023a). SENTIMENT ANALYSIS OF OMNIBUS LAW USING SUPPORT VECTOR MACHINE (SVM) WITH LINEAR KERNEL. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 17(4), 2197–2206. <https://doi.org/10.30598/barekengvol17iss4pp2197-2206>

- Makhtum, A. R., & Muhajir, M. (2023b). SENTIMENT ANALYSIS OF OMNIBUS LAW USING SUPPORT VECTOR MACHINE (SVM) WITH LINEAR KERNEL. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 17(4), 2197–2206. <https://doi.org/10.30598/barekengvol17iss4pp2197-2206>
- Makhtum, A. R., & Muhajir, M. (2023c). SENTIMENT ANALYSIS OF OMNIBUS LAW USING SUPPORT VECTOR MACHINE (SVM) WITH LINEAR KERNEL. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 17(4), 2197–2206. <https://doi.org/10.30598/barekengvol17iss4pp2197-2206>
- Makhtum, A. R., & Muhajir, M. (2023d). SENTIMENT ANALYSIS OF OMNIBUS LAW USING SUPPORT VECTOR MACHINE (SVM) WITH LINEAR KERNEL. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 17(4), 2197–2206. <https://doi.org/10.30598/barekengvol17iss4pp2197-2206>
- Matheos Sarimole, F., & Ihsan, A. N. (2023). ANALISIS SENTIMEN TWITTER TERHADAP UU CIPTA KERJA DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE TWITTER SENTIMENT ANALYSIS OF THE CIPTA KERJA LAW USING THE NAÏVE BAYES ALGORITHM AND SUPPORT VECTOR MACHINE. *Journal of Information Technology and Computer Science (INTECOMS)*, 6(2).
- Mustikasari, D., Widaningrum, I., Arifin, R., Henggal, W., & Putri, E. (2021). *Comparison of Effectiveness of Stemming Algorithms in Indonesian Documents*.
- Mutmainah, S., Khairunnas, & Khairunnisa. (2024). Metode Deep Learning LSTM dalam Analisis Sentimen Aplikasi PeduliLindungi. *Scientific : Journal of Computer Science and Informatics*, 1(1), 9–19. <https://doi.org/10.34304/scientific.v1i1.231>
- Ndruru, A. (2022a). Analisis Sentimen UU Cipta Kerja Melalui Omnibus Law Menggunakan Naive Bayes Classifier (NBC) Dan Support Vector Machine (SVM). *Pelita Informatika : Informasi dan Informatika*, 10(3).

- Ndruru, A. (2022b). Analisis Sentimen UU Cipta Kerja Melalui Omnibus Law Menggunakan Naive Bayes Classifier (NBC) Dan Support Vector Machine (SVM). *Pelita Informatika : Informasi dan Informatika*, 10(3).
- Pane, S. F., Owen, A., & Prianto, C. (2021a). Analisis Sentimen UU Omnibus Law pada Twitter Menggunakan Metode Support Vector Machine. *Jurnal Telekomunikasi dan Komputer*, 11(2), 130. <https://doi.org/10.22441/incomtech.v11i2.10874>
- Pane, S. F., Owen, A., & Prianto, C. (2021b). Analisis Sentimen UU Omnibus Law pada Twitter Menggunakan Metode Support Vector Machine. *Jurnal Telekomunikasi dan Komputer*, 11(2), 130. <https://doi.org/10.22441/incomtech.v11i2.10874>
- Prana, T., Sukma, W., & Pribadi, M. R. (2024). Analisis Sentimen Review Pengguna Viu pada Play Store dengan Algoritma Random Forest. *Journal of Software Engineering and Computational Intelligence (JSECI)*, 2(1).
- Pratama, E. B., & Marjun, L. A. (2022). Analisis Pemodelan Diagram Uml Pada Rancang Bangun Sistem Informasi Kepegawaian Yang Dikembangkan Dengan Model Waterfall. *Jurnal Teknik Informatika Kaputama (JTIK)*, 6(2), 725-736.
- Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023a). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 153-160. <https://doi.org/10.57152/malcom.v3i2.897>
- Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023b). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 153-160. <https://doi.org/10.57152/malcom.v3i2.897>
- Sabrani, A., Wedashwara, I. W. G. P. W., & Bimantoro, F. (2020). METODE MULTINOMIAL NAÏVE BAYES UNTUK KLASIFIKASI ARTIKEL

- ONLINE TENTANG GEMPA DI INDONESIA (Multinomial Naïve Bayes Method for Classification of Online Article About Earthquake in Indonesia).
- Sabrina, T. S., Azhar, Y., & Aditya, C. S. K. (2022). *Analisis Sentimen Tweet Tentang UU Cipta Kerja Menggunakan Algoritma SVM Berbasis PSO*. *Jurnal Informatika Sunan Kalijaga* (Vol. 7).
- Sandryan, M. K., Rahayudi, B., & Ratnawati, D. E. (2021a). *Analisis Sentimen Pada Media Sosial Twitter Terhadap Undang-Undang Cipta Kerja Menggunakan Algoritma Backpropagation dan Term Frequency-Inverse Document Frequency* (Vol. 5).
- Sandryan, M. K., Rahayudi, B., & Ratnawati, D. E. (2021b). *Analisis Sentimen Pada Media Sosial Twitter Terhadap Undang-Undang Cipta Kerja Menggunakan Algoritma Backpropagation dan Term Frequency-Inverse Document Frequency* (Vol. 5). Diambil dari <http://j-ptiik.ub.ac.id>
- Sarimole, F. M., & Ihsan, A. N. (2023). ANALISIS SENTIMEN TWITTER TERHADAP UU CIPTA KERJA DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE TWITTER SENTIMENT ANALYSIS OF THE CIPTA KERJA LAW USING THE NAÏVE BAYES ALGORITHM AND SUPPORT VECTOR MACHINE. *Journal of Information Technology and Computer Science (INTECOMS)*, 6(2).
- Sukma, E. A., Hidayanto, A. N., Pandesenda, A. I., Yahya, A. N., Widharto, P., & Rahardja, U. (2020a). Sentiment Analysis of the New Indonesian Government Policy (Omnibus Law) on Social Media Twitter. Dalam *Proceedings - 2nd International Conference on Informatics, Multimedia, Cyber, and Information System, ICIMCIS 2020* (hlm. 153–158). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICIMCIS51567.2020.9354287>
- Sukma, E. A., Hidayanto, A. N., Pandesenda, A. I., Yahya, A. N., Widharto, P., & Rahardja, U. (2020b). Sentiment Analysis of the New Indonesian Government Policy (Omnibus Law) on Social Media Twitter. Dalam *Proceedings - 2nd International Conference on Informatics, Multimedia, Cyber, and Information*

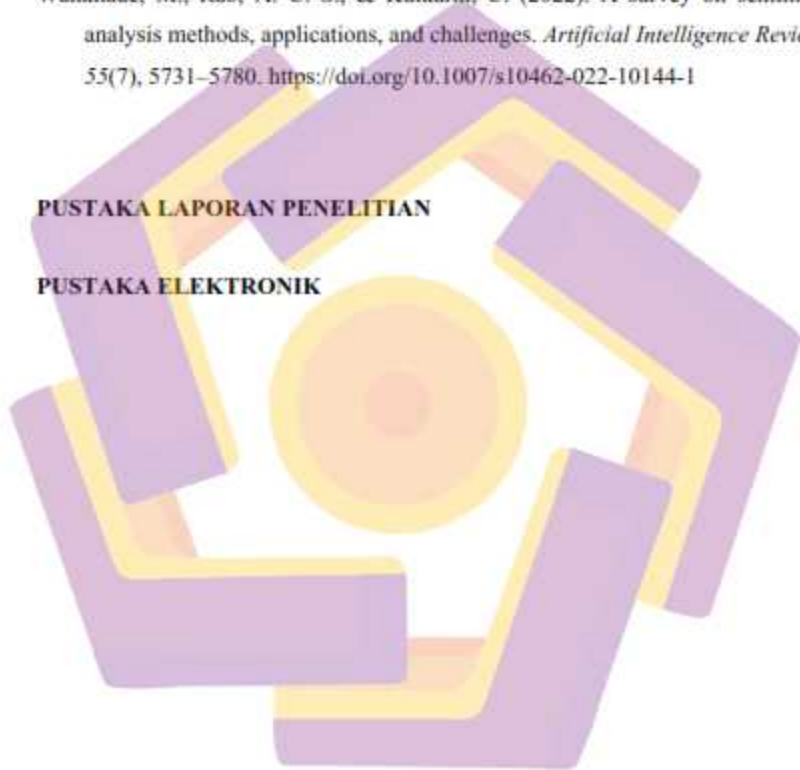
System, ICIMCIS 2020 (hlm. 153–158). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICIMCIS51567.2020.9354287>

Wang, Y., Guo, J., Yuan, C., & Li, B. (2022). Sentiment Analysis of Twitter Data. *Applied Sciences (Switzerland)*, 12(22). <https://doi.org/10.3390/app122211775>

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>

PUSTAKA LAPORAN PENELITIAN

PUSTAKA ELEKTRONIK



LAMPIRAN

Dataset Tweet dan Slang

<https://www.kaggle.com/datasets/arifdwisvafutra/dataset-tweet-omnibus-law-and-indonesian-slang>

