Indonesian Journal of Multidisciplinary on Social and Technology
Homepage: https://journal.ilmudata.co.id/index.php/ijmst
Vol. 5 No. 1 (2026) pp: 102-116
P-ISSN: 2986-6790, e-ISSN: 2986-6782

Perbandingan Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest dalam Analisis Sentimen Komentar Youtube terhadap UU TNI

Zikri Hanif, Andreo Yudertha
Prodi Sistem Informasi, Fakultas Sains Dan Teknologi, Universitas Islam Negeri Sultan Thaha Saifudin Jambi
zikrihanif2013@gmail.com, andreo@uinjambi.ac.id*

Abstrak

Platform YouTube menjadi tempat perdebatan publik yang dapat memfasilitasi penggunanya dalam mengutarakan pandangan secara bebas. Undang-Undang Tentara Nasional Indonesia (UU TNI) yang ditetapkan pada tahun 2025 menimbulkan perdebatan publik secara luas, terutama di platform media sosial YouTube. Isu utama dalam penelitian ini adalah bagaimana mengidentifikasi sentimen publik terhadap revisi UU TNI secara akurat serta menentukan algoritma machine learning yang memiliki performa terbaik dalam mengklasifikasikan opini publik berbahasa Indonesia di media sosial. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap revisi UU TNI serta mengkomparasikan kinerja algoritma yang paling baik. Metode yang digunakan meliputi pengumpulan data sebanyak 14.898 komentar YouTube menggunakan kata kunci "UU TNI", "Revisi UU TNI", dan "Undang-Undang TNI". Selanjutnya dilakukan tahap preprocessing data dan pelabelan menggunakan tiga metode, yaitu pelabelan manual (ground truth), pelabelan otomatis menggunakan InSet Lexicon, dan pseudo labeling berbasis IndoBERT. Setelah itu, data diklasifikasikan menggunakan algoritma Naïve Bayes, Support Vector Machine (SVM), dan Random Forest. Hasil penelitian menunjukkan bahwa sentimen publik didominasi oleh sentimen netral dan negatif, sedangkan sentimen positif memiliki jumlah yang lebih sedikit. Hasil perbandingan ketiga metode pelabelan menunjukkan bahwa skenario pseudo labeling berbasis IndoBERT menghasilkan performa klasifikasi terbaik. Selain itu, algoritma Support Vector Machine (SVM) menunjukkan performa terbaik secara keseluruhan dengan memperoleh akurasi sebesar 82,19% dan Macro F1-Score sebesar 0,75 pada skenario pseudo labeling IndoBERT, lebih tinggi dibandingkan Naïve Bayes dan Random Forest.

Kata kunci: Sistem Informasi, Penjualan Karet, UPPB, Website, Waterfall.

Abstract

YouTube has become a platform for public debate, enabling users to express their opinions freely. The revision of the Indonesian National Armed Forces Law (TNI Law) enacted in 2025 has generated widespread public controversy, particularly on YouTube. This study addresses the challenge of accurately identifying public sentiment toward the revised TNI Law and determining the most effective machine learning algorithm for classifying Indonesian-language opinions on social media. The research aims to analyze public sentiment regarding the revision and compare the performance of different classification algorithms. A total of 14,898 YouTube comments were collected using the keywords "UU TNI," "Revisi UU TNI," and "Undang-Undang TNI." The data underwent preprocessing and labeling using three approaches: manual labeling (ground truth), automatic labeling with the InSet Lexicon, and IndoBERT-based pseudo labeling. Subsequently, sentiment classification was performed using Naïve Bayes, Support Vector Machine (SVM), and Random Forest algorithms. The findings indicate that public sentiment was predominantly neutral and negative, while positive sentiment accounted for a smaller proportion. Comparisons across the three labeling approaches revealed that IndoBERT-based pseudo labeling achieved the highest classification performance. Furthermore, Support Vector Machine demonstrated the best overall performance, achieving an accuracy of 82.19% and a Macro F1-Score of 0.75 under the IndoBERT pseudo-labeling scenario, outperforming both Naïve Bayes and Random Forest. These findings highlight the effectiveness of combining contextual pseudo labeling with machine learning techniques for analyzing Indonesian public opinion.

Keywords: Sentiment Analysis, TNI Law Revision, YouTube, Machine Learning, IndoBERT, Support Vector Machine

1. Pendahuluan

Kemajuan sistem informasi dan internet menjadikan media sosial sebagai sarana diskusi, penyebaran informasi, dan pembentukan opini publik terkait permasalahan sosiopolitik [1]. Salah satu platform yang berpengaruh adalah YouTube, yang kini menjadi wadah portal berita nasional dan podcast dalam menyajikan informasi terkait

kebijakan pemerintah [2]. Kolom komentar YouTube menjadi sumber data teks yang kaya untuk menganalisis kecenderungan opini masyarakat melalui diskusi dan pendapat yang disampaikan pengguna.

Fenomena ini terlihat dalam perdebatan publik mengenai Undang-Undang TNI Tahun 2025. Pasal 47 ayat (2) Undang-Undang Nomor 3 Tahun 2025 mengatur kesempatan bagi anggota TNI menduduki jabatan sipil setelah tidak lagi berstatus aktif [3]. Kebijakan ini menimbulkan beragam tanggapan. Sebagian pihak menilai revisi tersebut dapat memperkuat sinergi antar lembaga negara, sedangkan pihak lain mengkhawatirkan dampaknya terhadap hubungan sipil dan militer di Indonesia [4]. Perdebatan ini ramai dibahas dalam kolom komentar video berita YouTube dengan opini yang mencerminkan sentimen positif, negatif, dan netral.

Dalam situasi demokrasi, opini publik di ruang digital berfungsi sebagai bentuk kontrol sosial, di mana masyarakat dapat memberikan dukungan, kritik, dan masukan terhadap kebijakan yang dikeluarkan pemerintah [5]. Pemilihan UU TNI sebagai objek penelitian didasarkan pada sifatnya sebagai isu strategis nasional yang menyangkut hubungan sipil dan militer di Indonesia. Topik ini penting karena masih sedikit penelitian yang membahas analisis sentimen publik terhadap isu kebijakan strategis pemerintah, khususnya yang berkaitan dengan peran militer di ranah sipil. Analisis sentimen terhadap pendapat masyarakat di media digital diharapkan dapat menjelaskan mengenai persepsi publik tentang kebijakan tersebut secara sistematis dan berbasis data.

Untuk mengurai dan memahami volume opini publik yang sangat besar, heterogen, dan tidak terstruktur dari media sosial, pendekatan analisis sentimen (*sentiment analysis*) menjadi metode yang paling esensial untuk diimplementasikan. Analisis sentimen adalah komponen dari *text data mining* yang digunakan untuk mengelompokkan opini publik ke dalam kategori positif, negatif, atau netral secara otomatis. Metode ini dilaksanakan dengan memanfaatkan klasifikasi yang dilakukan secara cepat, konsisten, dan objektif dibandingkan pengamatan manual [6]. Penelitian ini menggunakan tiga metode pengelompokan yang diterapkan analisis sentimen, yaitu *Naïve Bayes, Support Vector Machine (SVM),* dan *Random Forest*.

Naïve Bayes dikenal sebagai metode pengelompokan berbasis probabilitas yang sederhana namun efisien. Metode ini memiliki kemampuan melakukan proses klasifikasi dengan waktu komputasi yang relatif cepat serta mampu menghasilkan performa yang baik meskipun menggunakan data latih yang terbatas. Selain itu, *Naïve Bayes* cukup tahan terhadap data berdimensi tinggi dan *noise*, sehingga sering digunakan dalam analisis teks pada media sosial. Hasil riset yang dilakukan oleh [7] ,menunjukkan bahwa algoritma *Naïve Bayes Classifier* mampu memperoleh nilai akurasi mencapai 84% dalam menganalisis sentimen aplikasi brimo menggunakan data dari google play store.

Selanjutnya *Support Vector Machine (SVM),* yang mendapat kemampuan generalisasi yang baik dalam menangani data berdimensi tinggi dan kompleks. *SVM* bekerja dengan mencari *hyperplane* optimal untuk memisahkan setiap kelas sentimen sehingga menghasilkan proses klasifikasi yang lebih akurat. Karakteristik tersebut menjadikan *SVM* kerap digunakan dalam pengembangan model analisis sentimen berbasis ekstraksi fitur *TF-IDF*. Penelitian yang dilakukan oleh [8] menunjukkan bahwa algoritma *SVM* memperoleh tingkat akurasi 85,1% dalam menganalisis sentimen masyarakat Indonesia terhadap Pilpres 2024.

Sementara itu, *Random Forest* termasuk algoritma *ensemble learning* yang menyatukan banyak *decision tree* guna menaikkan stabilitas dan akurasi prediksi. Metode ini dikenal mampu menangani data yang kompleks maupun mengandung *noise*. Penelitian [9] menunjukkan bahwa algoritma *Random Forest* memperoleh tingkat akurasi 70% pada analisis sentimen makan bergizi gratis.

Meskipun ketiga metode menunjukkan performa baik dalam penelitian sebelumnya, hasil klasifikasi bergantung pada karakter data dan topik yang dianalisis. Oleh karena itu, penelitian ini membandingkan performa Naïve Bayes, Support Vector Machine (SVM), dan Random Forest dalam mengelompokkan sentimen komentar publik YouTube terkait revisi UU TNI. Penelitian ini diharapkan dapat menentukan algoritma paling efektif, berkontribusi pada pengembangan machine learning, serta memberikan pemahaman mengenai opini publik terhadap kebijakan tersebut.

Berdasarkan uraian tersebut, penelitian ini mengangkat topik "Perbandingan Algoritma *Naïve Bayes, Support Vector Machine,* dan *Random Forest* dalam Analisis Sentimen Komentar YouTube terhadap UU TNI".

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

2. Metodelogi Penelitian

Metodologi penelitian merupakan susunan tahapan yang dikerjakan dalam pelaksanaan penelitian, dimulai dari proses pengambilan data hingga tahap penilaian hasil. Setiap tahapan disusun secara sistematis untuk mendukung tercapainya tujuan penelitian. Penelitian ini menggunakan metode kuantitatif karena proses analisis dilakukan berdasarkan data yang diolah secara komputasional dan dievaluasi menggunakan ukuran numerik[10]. Data komentar YouTube terlebih dahulu diproses menjadi representasi vektor melalui teknik ekstraksi fitur, kemudian dianalisis menggunakan algoritma *machine learning*. Berdasarkan tujuannya, penelitian ini termasuk dalam penelitian komparatif karena berfokus pada perbandingan performa beberapa algoritma klasifikasi, yaitu *Naïve Bayes, Support Vector Machine (SVM),* dan *Random Forest*, dalam menyelesaikan permasalahan analisis sentimen pada topik Undang-Undang Tentara Nasional Indonesia (UU TNI).

Pemilihan metode komparatif digunakan untuk membandingkan kinerja setiap algoritma dalam mengelompokkan sentimen pada data komentar YouTube. Hasil perbandingan tersebut digunakan untuk mengidentifikasi algoritma yang menunjukkan performa paling baik dalam mengidentifikasi data berdasarkan kelas positif, negatif, dan netral. Ukuran evaluasi yang digunakan sebagai tolok ukur performa berupa akurasi (*accuracy*), ketepatan (*precision*), sensitivitas (*recall*), dan rata-rata harmonik (*F1-score*).

Untuk meningkatkan jumlah data berlabel dan mengurangi kebutuhan anotasi manual dalam skala besar, penelitian ini menerapkan pendekatan pelabelan kombinasi *(hybrid labeling approach*). Proses pelabelan data sentimen dilakukan melalui tiga skenario utama, yaitu pelabelan manual *(ground truth)*, pelabelan berbasis *InSet Lexicon*, dan *pseudo labeling* menggunakan model IndoBERT.

2.1. Tahap Penelitian

Tahap kegiatan awal penelitian yang meliputi identifikasi masalah, perumusan masalah, penentuan batasan penelitian, penyusunan tujuan penelitian, serta studi literatur yang berkaitan dengan analisis sentimen, *machine learning*, *text mining*, metode pelabelan data, dan analisis mengenai Undang-Undang Tentara Nasional Indonesia (UU TNI). Studi literatur bertujuan memperoleh dasar teoritis sebagai acuan untuk mendukung proses penelitian serta membantu dalam menentukan metode yang tepat dalam menyelesaikan topik penelitian.

2.2. Tahap Pengumpulan Data

Tahap pengumpulan data *(data collection)* dilakukan menggunakan teknik ekstraksi data *(scraping)* terhadap opini publik yang terdapat pada platform YouTube. Data penelitian diperoleh secara komputasional menggunakan bahasa pemrograman *Python* pada lingkungan *Google Colaboratory* dengan memanfaatkan layanan resmi YouTube Data API v3.

Pengumpulan data diawali dengan pencarian video yang berkaitan dengan topik penelitian menggunakan kata kunci seperti “Revisi UU TNI”, “RUU TNI”, dan “UU TNI Terbaru”. Pencarian dibatasi pada video-video yang paling relevan dengan topik penelitian untuk menjaga kualitas data yang diperoleh. Setelah memperoleh ID video, sistem mengekstraksi komentar tingkat atas (*Top Level Comment*) dari setiap video. Data yang berhasil dikumpulkan meliputi nama pengguna (*username*), isi komentar, waktu komentar (*timestamp*), tanggal rilis video, dan ID video. Seluruh data kemudian diekspor ke bentuk file *Comma-Separated Values (CSV)* agar mempermudah proses selanjutnya.

2.3. Tahap Pengolahan Data

2.3.1 Pra-pemrosesan Data (Text Preprocessing)

Komentar yang diperoleh melalui proses *scraping* belum dapat digunakan secara langsung karena masih mengandung berbagai unsur yang tidak diperlukan dalam proses analisis. Oleh sebab itu, dilakukan tahap *text preprocessing* untuk meningkatkan kualitas data dengan membersihkan serta menyamakan format teks sehingga siap digunakan pada proses klasifikasi melalui algoritma *machine learning[11]*. Tahap pra-pemrosesan yang dilakukan:

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

1. *Cleansing*, yaitu proses membersihkan data dengan menghilangkan unsur-unsur yang tidak digunakan, meliputi URL, angka, , *hashtag*, dan *mention*.
2. *Case Folding*, merupakan tahap menyamakan komentar dengan cara mengonversi seluruh huruf menjadi huruf kecil (*lowercase*).
3. *Tokenizing*, yaitu tahapan memisahkan setiap kalimat menjadi kumpulan token atau kata sehingga dapat diproses pada tahap berikutnya.
4. *Stopword Removal*, yaitu tahap menghilangkan kata yang sering muncul yang tidak memberikan informasi penting terhadap tahap klasifikasi sentimen dengan memanfaatkan pustaka Sastrawi dan daftar *stopword* tambahan.
5. *Stemming*, yaitu tahap mengubah kata yang menggunakan pustaka Sastrawi yang didukung oleh *custom dictionary* untuk mempertahankan kata-kata tertentu yang dianggap penting.

2.3.2. Pelabelan Data Dan Pembentukan Scenario Eksperimen

Setelah seluruh tahapan *text preprocessing* selesai, *dataset* diperiksa untuk mengidentifikasi dan menghapus data yang memiliki isi komentar yang sama sehingga tidak terjadi pengulangan data pada proses analisis[12]. Langkah ini dilakukan agar kualitas *dataset* tetap terjaga dan hasil evaluasi model lebih representatif. Selanjutnya, *dataset* disusun ke dalam tiga skenario eksperimen berdasarkan metode pelabelan yang diterapkan.

1. Skenario manual (*ground truth*) menggunakan *dataset* yang seluruh label sentimennya ditentukan melalui proses pelabelan oleh peneliti. Setiap komentar diklasifikasikan berdasarkan isi dan konteksnya ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. *Dataset* ini digunakan sebagai acuan dalam proses pelatihan serta evaluasi performa model klasifikasi. Penentuan label pada skenario manual dilakukan berdasarkan kriteria berikut: Sentimen Positif, Sentimen Negatif dan Sentimen Netral.
2. Skenario *Pseudo Labeling* IndoBERT menggunakan model IndoBERT untuk memberikan label otomatis pada data yang belum memiliki label. Untuk menjaga kualitas hasil pelabelan, hanya prediksi dengan *confidence score* $\geq 0{,}90$ yang digunakan sebagai pseudo label.
3. Skenario *InSet Lexicon* menggunakan kamus sentimen Bahasa Indonesia *InSet Lexicon* untuk memberikan label sentimen berdasarkan akumulasi skor polaritas kata pada setiap komentar.

Ketiga skenario tersebut digunakan secara terpisah pada proses pelatihan dan evaluasi model untuk mengetahui pengaruh metode pelabelan terhadap performa klasifikasi sentimen.

2.3.4. Pembobolan Data (Feature Extraction)

Data yang telah dilabel kemudian dirubah menjadi bentuk numerik dengan metode *Term Frequency-Inverse Document Frequency(TF-IDF)*. Pada tahap pembentukan fitur,. Nilai setiap fitur ditentukan berdasarkan seberapa sering kemunculan kata pada dokumen serta tingkat distribusi secara keseluruhan pada dokumen di *dataset*. *Output* dari proses tersebut dalam bentuk matriks fitur yang Kemudian digunakan sebagai data masukan pada tahap pelatihan dan pengujian model *algoritma Naïve Bayes, Support Vector Machine (SVM),* dan *Random Forest[13].*

2.3.5. Pelatihan Model dan Klasifikasi Sentimen

Matriks fitur yang dihasilkan melalui proses penilaian TF-IDF selanjutnya digunakan sebagai data masukan pada tahap klasifikasi. Penelitian ini membandingkan tiga algoritma *machine learning*, yaitu *Naïve Bayes, Support Vector Machine*, dan *Random Forest*. Masing-masing algoritma diterapkan pada setiap skenario *dataset* untuk menghasilkan model klasifikasi sentiment[14]. Selama proses pelatihan, setiap model mempelajari karakteristik fitur yang berasal dari hasil pembobotan TF-IDF sehingga dapat mengelompokkan komentar ke dalam kategori sentimen yang sesuai.

2.3.6 Pengujian Model (Validation)

Evaluasi model pada penelitian ini dilakukan menggunakan *Stratified 10-Fold Cross Validation.* Metode tersebut dipilih karena setiap pembagian data tetap mempertahankan proporsi masing-masing kelas sentimen, sehingga distribusi data pada setiap *fold* tetap seimbang. Dengan metode ini, hasil penilaian dapat menggambarkan performa model secara lebih konsisten serta meminimalkan pengaruh ketidakseimbangan data selama proses validasi.

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Pada setiap putaran pengujian, *dataset* dibagi menjadi sepuluh *fold* dengan proporsi kelas sentimen yang tetap dipertahankan pada setiap *fold*. Secara bergantian, satu *fold* digunakan sebagai data pengujian (*testing set*), sedangkan sembilan *fold* lainnya digunakan sebagai data pelatihan (*training set*)[15]. Proses tersebut diulang sebanyak sepuluh kali hingga seluruh *fold* memperoleh kesempatan satu kali sebagai data pengujian. Nilai rata-rata *accuracy, precision, recall,* dan *F1-score* dari seluruh putaran digunakan sebagai indikator performa akhir masing-masing algoritma klasifikasi.

2.4. Tahapan Evaluasi dan Analisis Hasil

Tahap evaluasi bertujuan menilai hasil klasifikasi yang diperoleh dari algoritma *Naïve Bayes, Support Vector Machine (SVM),* dan *Random Forest* pada setiap skenario pelabelan. Selain digunakan untuk membandingkan kemampuan masing-masing algoritma, tahap ini juga bertujuan mengidentifikasi perbedaan hasil klasifikasi yang dihasilkan oleh setiap metode pelabelan.

Penilaian performa Selain itu, *Confusion Matrix* dimanfaatkan untuk menggambarkan hubungan antara hasil prediksi dan label aktual, sehingga kesalahan prediksi pada setiap kategori sentimen dapat dianalisis secara lebih rinci. Berdasarkan hasil evaluasi tersebut, dilakukan identifikasi terhadap kombinasi algoritma dan metode pelabelan yang menghasilkan kinerja paling baik dalam analisis sentimen komentar YouTube mengenai revisi Undang-Undang Tentara Nasional Indonesia (UU TNI).

2.5. Tahap Visualisasi Data dan Pelaporan

Hasil penelitian disajikan dalam bentuk visualisasi berupa grafik perbandingan nilai *Accuracy, Precision, Recall, dan F1-Score* guna mempermudah analisis serta interpretasi kinerja setiap model pada skenario pengujian. Selain itu, *Confusion Matrix* digunakan untuk memberikan gambaran yang lebih rinci mengenai hasil klasifikasi dengan menampilkan distribusi prediksi model terhadap setiap kelas sentimen.

Karakteristik *dataset* juga divisualisasikan menggunakan *WordCloud* untuk menunjukkan kata-kata yang paling sering muncul pada komentar pengguna YouTube terkait UU TNI. Tahap akhir penelitian dilakukan dengan menyusun laporan skripsi yang mendokumentasikan seluruh rangkaian penelitian, mulai dari proses pengumpulan dan pengolahan data, hasil pengujian, analisis, hingga penarikan kesimpulan.

3. Analisis dan Hasil

3.1. Analisis Pendahuluan

Analisis Memiliki peran penting dalam penelitian karena pada tahap ini dilakukan pengkajian secara sistematis terhadap permasalahan yang diteliti. Melalui proses analisis, peneliti dapat memperoleh pemahaman yang lebih jelas mengenai objek penelitian serta menarik kesimpulan berdasarkan data yang telah diperoleh. Selain itu, analisis yang dilakukan secara tepat akan membantu dalam menentukan langkah-langkah penelitian secara terarah, mulai dari proses pengolahan data hingga interpretasi hasil yang diperoleh.

3.1.1. Analisis Studi Kasus

Tahap ini meliputi analisis opini publik terkait UU TNI berdasarkan komentar pengguna pada berbagai video YouTube. Platform ini dipilih karena banyak dimanfaatkan masyarakat untuk memperoleh informasi dan menyampaikan pendapat, baik berupa dukungan, kritik, maupun pandangan netral.

Berdasarkan pengamatan, komentar terkait UU TNI menunjukkan beragam pandangan dan intensitas diskusi yang cukup tinggi. Namun, analisis komentar tersebut secara sistematis menggunakan metode machine learning masih jarang dilakukan. Oleh karena itu, penelitian ini berfokus pada analisis sentimen publik terhadap kebijakan UU TNI secara terstruktur dan berbasis data.

3.1.2. Pengumpulan Data

Pengambilan data pada penelitian ini dilakukan menggunakan proses *scraping* dari platform YouTube menggunakan bahasa pemrograman *Python* dengan memanfaatkan YouTube Data API v3. Proses pengambilan

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

data dilakukan secara otomatis dengan mencari video berdasarkan beberapa kata kunci, yaitu “Revisi UU TNI”, “UU TNI Terbaru”, dan “RUU TNI”, sehingga data yang diperoleh relevan dengan topik penelitian serta mencerminkan opini publik yang berkembang di masyarakat. Pengambilan data dibatasi pada periode 1 Maret 2025 hingga 31 Maret 2026, serta didapatkan 89 video yang relevan dengan kata kunci yang digunakan. Dari setiap video yang diperoleh, dilakukan pengambilan komentar pengguna sebagai data penelitian. Berdasarkan proses *scraping* yang dilakukan, diperoleh sebanyak 14.898 data komentar (*raw data*) yang selanjutnya digunakan sebagai *dataset* dalam penelitian ini.

3.1.3. Data Komentar

Data yang diperoleh dari proses *scraping* berupa data komentar pengguna YouTube yang berkaitan dengan topik revisi UU TNI. *Dataset* yang dihasilkan terdiri dari beberapa atribut, yaitu tanggal video rilis, judul video, tanggal komentar, *username*, dan isi komentar berjumlah 14.898 komentar.

Tabel 3.1 Data Komentar

index	Tanggal_Video_Rilis	Judul_Video	Tanggal_Komentar	Username	Komentar
1	2025-03-19	Revisi UU TNI: Apa dampaknya untuk masyarakat sipil dan mengapa buru-buru disahkan?	2026-03-27	@SamsungBekas-hc9kk	DPR NYA AMBURADUL , APAN TNI DUDUKI KENERJA SIPIL , TNI BENTENG NEGARA DAN MENGAYOMI RAKYAT SEMACAM POLISI
2	2025-03-19	Revisi UU TNI: Apa dampaknya untuk masyarakat sipil dan mengapa buru-buru disahkan?	2025-06-02	@nikearela7033	Yang nganggur masih banyak oy,kasihlah yang laen kesempatan
3	2025-03-17	DPR Klaim Revisi UU TNI untuk Membatasi, Bukan untuk Memperluas Jabatan Prajurit di Sektor Sipil	2025-03-17	@andreassudarman8283	BikinnYa aja sembunyi2... Ngeluh banyak jendral nganggur.. 😅😅😅😅😅...
4	2025-03-20	BATALKAN REVISI UU TNI	2025-06-03	@dewasutedja2261	Apa hasil kerja dari TNI ???? coba siapa yang bisa menjawabnya.....dan siapa penghianat reformasi coba juga jawab ? masih ingat kasusHAM trisakti, kudatuli, tanjung periok, Marsinah ?? siapa yang bisa menjawab ?
5	2026-03-15	Remiliterisasi dan Judicial Review UU TNI (Zainal Arifin & Andrie Yunus)	2026-03-19	@ONEpisssss	sudah lupa anak anak bangsa yg berkarier di militer tugas dan fungsinya... yaitu menjaga keutuhan NKRI dan persatuan Indonesia.. mereka sibuk nyari side job yg dipikirkan cuan.. main proyek, main tambang, lupa kalau kalian dititipkan jabatan buat menjaga NKRI.. jgn musuhi rakyat

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

3.2. Preprocessing Data

Tahap *preprocessing* data berfungsi untuk meningkatkan kualitas data sebelum diproses lebih lanjut, proses analisis sentimen bertujuan meningkatkan kualitas data teks sebelum digunakan pada proses klasifikasi. Komentar yang diperoleh dari hasil *scraping* masih mengandung berbagai unsur yang dapat memengaruhi proses analisis, seperti URL, karakter khusus, simbol, penggunaan huruf kapital, serta kata-kata yang tidak mengikuti kaidah bahasa baku. Oleh karena itu, dilakukan serangkaian tahapan *preprocessing* yang meliputi *cleansing*, normalisasi, *stopword removal*, *tokenizing*, dan *stemming*. Melalui tahapan tersebut, Data diproses sehingga menjadi lebih bersih, konsisten, agar siap digunakan oleh algoritma *machine* learning.

3.2.1. Normalisasi

Normalisasi merupakan tahap preprocessing untuk menyeragamkan teks agar mudah dianalisis. Proses ini meliputi cleansing untuk menghapus elemen yang tidak relevan, seperti URL, mention, simbol, angka, dan karakter khusus; case folding untuk mengubah huruf menjadi lowercase; serta penggantian kata tidak baku, singkatan, dan bahasa gaul menjadi kata baku menggunakan kamus normalisasi agar teks lebih seragam sebelum diproses oleh algoritma machine learning.

Tabel 3.2 Normalisasi

no	Komentar	Teks_bersih
1	Mungkin..hanya ..TNI ..yg msh menyala sampai saat ini....hidup prabiwao..hidup rakyat....!	mungkinhanya tni yang masih menyala sampai saat inihidup prabiwaohidup rakyat
2	DPR NYA AMBURADUL , APAN TNI DUDUKI KENERJA SIPIL , TNI BENTENG NEGARA DAN MENGAYOMI RAKYAT SEMACAM POLISI	dpr nya amburadul apan tni duduki kenerja sipil tni benteng negara dan mengayomi rakyat semacam polisi
3	BAGUS DONG TNI DI LIBATKAN , KOLABORASI , SAYA SETUJU , BERSINERGI , KOMPAK, SOLID , NEGARA LEBIH AMAN , LEBIH KEPADA BERSIH , MANUSIAWI , BERWIBAWA❤️ ❤️ ❤️ ❤️ GBU ALL TNI & POLRI	bagus dong tni di libatkan kolaborasi saya setuju bersinergi kompak solid negara lebih aman lebih kepada bersih manusiawi berwibawa gbu all tni polri

3.2.2. Stopword Removal

Stopword removal merupakan tahapan *preprocessing* yang bertujuan menghilangkan kata-kata dengan frekuensi kemunculan tinggi tetapi memiliki kontribusi yang sangat kecil terhadap proses klasifikasi sentimen. Kata-kata seperti "dan", "yang", dan "di" dihapus karena tidak memberikan informasi yang relevan dalam menentukan kelas sentimen. Pada penelitian ini, proses *stopword removal* dilakukan menggunakan pustaka Sastrawi yang dipadukan dengan daftar *custom stopword*. Daftar tersebut mencakup kata-kata tidak baku atau bahasa percakapan, seperti "yg" dan "aja", yang sering muncul pada komentar pengguna, namun tidak memiliki nilai sentimen yang signifikan.

Tabel 3.3 Stopword Removal

No	Teks_bersih	*Teks_stopword*
1	mungkin yang bingung para koruptor kali	mungkin bingung koruptor kali
2	tolong siapa yang bikin negara dan rakyatnya miskin gara gara korupsi uang negara dan uang rakyat	siapa bikin negara rakyatnya miskin gara gara korupsi uang negara uang rakyat
3	ngerih kalo sudah jdi seperti myanmar masyarakat yang melawan hanya sampah bagi mereka	ngerih kalo jdi myanmar masyarakat melawan sampah mereka

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

3.2.3. Tokenizing

Tokenizing merupakan proses memecah kalimat menjadi kata-kata atau *token* berdasarkan spasi sehingga memudahkan dalam proses analisis. Pada penelitian ini, *tokenizing* dilakukan dengan memisahkan setiap kata dalam kalimat menggunakan fungsi sederhana pada *Python*. Hasil dari proses ini berupa daftar kata (*token*) yang akan digunakan pada tahap selanjutnya, sehingga mempermudah sistem dalam mengenali pola kata pada data teks.

Tabel 3.4 Tokenizing

No	Teks_stopword	Teks_Token
1	dpr nya amburadul apan duduki kenerja sipil benteng negara mengayomi rakyat semacam polisi	dpr,nya,amburadul,apan,,duduki,kenerja,sipil,,benteng,negara,mengayomi,rakyat,semacam,polisi
2	males dengerin omongan lembaga paling kebal hukum anti kritik indonesia	males,dengerin,omongan,lembaga,paling,kebal,hukum,anti,kritik,indonesia

3.2.4. Stemming

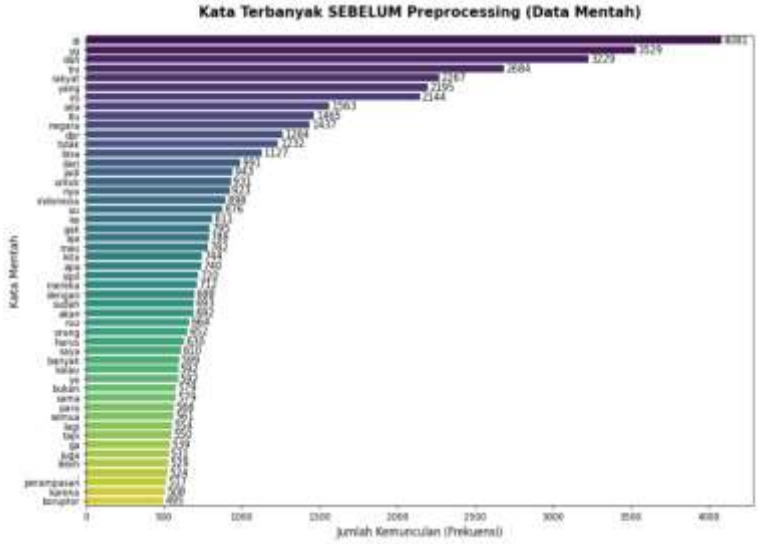
Stemming merupakan proses mengubah kata menjadi bentuk dasar untuk mengurangi keragaman kata dalam teks menggunakan pustaka Sastrawi. Proses ini juga memanfaatkan custom dictionary untuk mempertahankan makna kata tertentu, terutama yang berkaitan dengan sentimen. Contohnya, kata “menyetujui” dan “disetujui” menjadi “setuju”, serta “perampasan” menjadi “rampas”. Setelah tokenisasi, setiap kata diubah ke bentuk dasar dan digabung kembali menjadi kalimat. Hasilnya berupa teks yang lebih sederhana dan konsisten sehingga memudahkan proses pembobotan dan klasifikasi sentimen.

Tabel 3.5 Stemming

No	*Teks_token*	*Teks_stemming*
1	[demo, jaman, sekarang, kebanyakan, bayaran]	demo jaman sekarang banyak bayar
2	prabowo,sdg,membangun,pemerintahan,diktator,militer,stop,militerisasi	prabowo sdg bangun perintah diktator militer stop militer

3.2.5. Analisis Frekuensi Kata

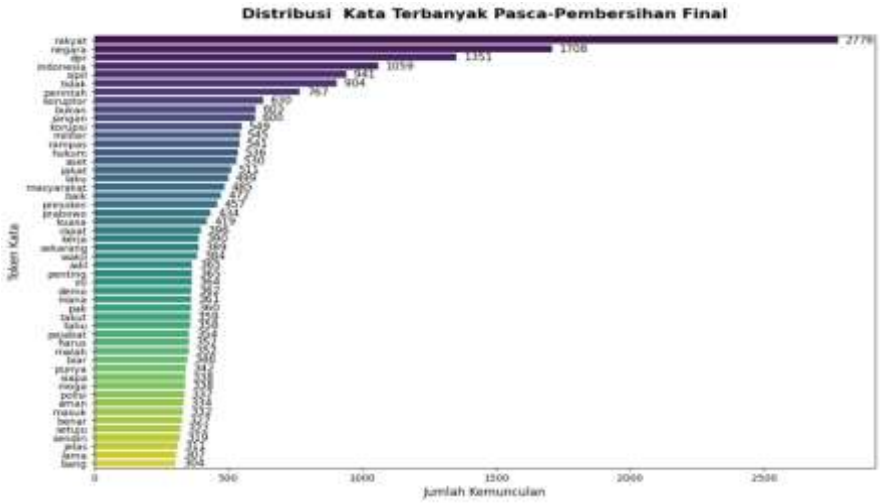
Analisis Jumlah kemunculan kata dilakukan untuk Menentukan kata-kata yang paling sering muncul pada *dataset* sebelum dan setelah melalui tahap *preprocessing*. Hasil analisis pada data sebelum *preprocessing* menunjukkan bahwa komentar masih didominasi oleh kata-kata umum serta elemen yang kurang memberikan informasi terhadap proses analisis sentimen. Kondisi tersebut menunjukkan bahwa data masih memerlukan proses pembersihan agar kata-kata yang lebih relevan dapat diidentifikasi pada tahapan berikutnya.



Gambar 3.1 *Frekunsi data sebelum* Preprocessing

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Setelah melalui tahap *preprocessing*, kata-kata yang terdapat pada *dataset* menjadi lebih representatif terhadap isi komentar dan topik penelitian. Berkurangnya kata-kata umum serta elemen yang tidak relevan menunjukkan bahwa tahapan *preprocessing* mampu meminimalkan *noise* pada data. Dengan demikian, kualitas *dataset* menjadi lebih baik sehingga dapat mendukung proses ekstraksi fitur dan klasifikasi sentimen secara lebih optimal.



Gambar 3.2 *Frekunsi data sebelum* Preprocessing

3.3. Pelebelan Data

Proses pelabelan sentimen dilakukan untuk Menggolongkan komentar pengguna YouTube ke dalam kategori sentimen,. Pelabelan dilakukan menggunakan tiga pendekatan yang berbeda untuk membentuk tiga skenario eksperimen, yaitu pelabelan manual *(ground truth), pseudo labeling* menggunakan IndoBERT, dan pelabelan berbasis kamus sentimen *InSet Lexicon*.

Sebelum proses pelabelan dilakukan, *dataset* hasil *scraping* terlebih dahulu melalui tahap pembersihan data dan penghapusan komentar duplikat. Proses ini bertujuan untuk mengurangi bias yang dapat memengaruhi hasil klasifikasi serta meningkatkan kualitas *dataset* yang digunakan pada penelitian.

3.3.1. Skenario Manual (*Ground Truth*)

Pelabelan manual dilakukan berdasarkan kriteria yang telah dijelaskan. Pendekatan ini dikenal sebagai *ground truth* karena anotator manusia memiliki kemampuan memahami konteks kalimat, nuansa emosi, sarkasme, serta penggunaan bahasa tidak baku yang sering kali sulit dikenali oleh metode pelabelan otomatis. Setelah proses penghapusan data duplikat dilakukan, diperoleh sebanyak 1.394 komentar yang digunakan pada skenario *ground truth*.

3.3.2. *Skenario Pseudo Labeling* IndoBERT

Pada skenario kedua, pelabelan dilakukan menggunakan model IndoBERT. Model ini digunakan untuk memprediksi label sentimen pada komentar yang belum memiliki label. Untuk menjaga kualitas hasil pelabelan otomatis, hanya prediksi dengan *confidence score* $\geq 0{,}90$ yang digunakan sebagai pseudo label. Setelah proses penghapusan data duplikat dilakukan, diperoleh sebanyak 8.614 komentar yang digunakan pada skenario IndoBERT.

3.3.3. Skenario *InSet Lexicon*

Setelah proses pembersihan data dan penghapusan duplikat dilakukan, diperoleh sebanyak 10.936 komentar yang digunakan pada skenario *InSet Lexicon.*. Berdasarkan hasil pelabelan menggunakan *InSet Lexicon*, sentimen negatif mendominasi *dataset* dengan proporsi yang jauh lebih besar dibandingkan sentimen positif maupun netral.

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Kondisi ini menunjukkan bahwa metode pelabelan berbasis leksikon cenderung mengidentifikasi lebih banyak komentar yang mengandung kata-kata bernuansa negatif.

3.3.4. Ringkasan *Dataset* Penelitian

Ketiga dataset tersebut digunakan secara terpisah dalam proses pembobotan TF-IDF, pelatihan model, dan evaluasi performa algoritma *Naïve Bayes, Support Vector Machine (SVM), dan Random Forest*. Pendekatan ini dilakukan untuk menganalisis pengaruh metode pelabelan terhadap performa klasifikasi sentimen pada komentar YouTube terkait UU TNI.

3.4. Pembobotan TF-IDF

Tahap selanjutnya adalah pembobotan kata menggunakan Term Frequency–Inverse Document Frequency (TF-IDF) untuk mengubah teks hasil preprocessing menjadi representasi numerik. Pembobotan ini didasarkan pada frekuensi kemunculan kata dalam dokumen dan penyebarannya pada seluruh dataset. Kata yang lebih spesifik terhadap suatu dokumen memperoleh bobot lebih tinggi, sedangkan kata yang sering muncul di banyak dokumen cenderung memiliki bobot lebih rendah.

Pembobotan TF-IDF dilakukan secara terpisah pada setiap skenario pelabelan, yaitu Ground Truth, IndoBERT, dan InSet Lexicon, agar representasi fitur sesuai dengan karakteristik masing-masing dataset.

Hasil pembobotan berupa matriks fitur numerik yang digunakan sebagai masukan untuk pelatihan dan pengujian model klasifikasi sentimen menggunakan algoritma Naïve Bayes, Support Vector Machine (SVM), dan Random Forest. Matriks tersebut membantu setiap algoritma mempelajari pola dalam data teks untuk melakukan klasifikasi sentimen.

Tabel 3.6 Data Bobot TF-IDF

No	*Word*	*Attribute Name*	*Total Occurrences*	*Document Occurrences*	*Average TF-IDF*	*Maximum TF-IDF*
1	rakyat	rakyat	2930	2026	0.0320	0.6691
2	enggak	enggak	2209	1761	0.0258	1.0000
3	negara	negara	1767	1276	0.0205	0.6929
4	jadi	jadi	1794	1396	0.0196	1.0000
5	indonesia	indonesia	1162	929	0.0153	0.5477
6	apa	apa	1095	933	0.0149	0.6044
7	orang	orang	1124	822	0.0145	0.6971
8	sama	sama	969	820	0.0134	0.7265
9	sipil	sipil	999	714	0.0131	0.7428
10	jabat	jabat	924	704	0.0120	0.6685

3.5. Stratified 10-Fold Cross Validation

Tahap selanjutnya adalah mengevaluasi kinerja model menggunakan metode Stratified 10-Fold Cross Validation yang mempertahankan proporsi kelas sentimen pada setiap fold. Data dibagi menjadi sepuluh fold, dengan sembilan fold digunakan untuk pelatihan (training set) dan satu fold untuk pengujian (testing set). Proses ini

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

diulang sepuluh kali hingga setiap fold menjadi data pengujian. Hasil evaluasi setiap iterasi kemudian dirata-ratakan untuk memperoleh nilai performa akhir masing-masing algoritma.

Evaluasi dilakukan menggunakan metrik Accuracy, Precision, Recall, dan F1-Score untuk mengukur kemampuan model dalam mengklasifikasikan sentimen. Selain itu, Confusion Matrix digunakan untuk membandingkan hasil prediksi dengan label aktual serta menganalisis kesalahan klasifikasi pada setiap kategori sentimen.

Tabel 3.7 K-fold Ground Truth

Fold	Training	Testing	Train Negatif	Train Netral	Train Positif	Test Negatif	Test Netral	Test Positif
1	918	103	373	407	138	41	46	16
2	919	102	373	407	139	41	46	15
3	919	102	373	407	139	41	46	15
4	919	102	372	408	139	42	45	15
5	919	102	372	408	139	42	45	15
6	919	102	372	408	139	42	45	15
7	919	102	372	408	139	42	45	16
8	919	102	373	408	138	41	45	16
9	919	102	373	408	138	41	45	16
10	919	102	373	408	138	41	45	16

3.6. Performa Model Algoritma

Pada tahap pengujian, algoritma Naïve Bayes, Support Vector Machine (SVM), dan Random Forest digunakan untuk mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu negatif, netral, dan positif. Evaluasi dilakukan menggunakan Stratified 10-Fold Cross Validation untuk mempertahankan proporsi kelas sentimen pada setiap fold dan memperoleh hasil evaluasi yang lebih konsisten.

Pengujian dilakukan pada tiga skenario dataset, yaitu Ground Truth (manual), IndoBERT, dan InSet Lexicon. Kinerja setiap algoritma diukur menggunakan metrik Accuracy, Precision, Recall, dan F1-Score untuk membandingkan kemampuan klasifikasi pada setiap skenario.

3.6.1. Hasil Pengujian Skenario Ground Truth

Skenario *Ground Truth* menggunakan *dataset* yang seluruh labelnya diberikan secara manual oleh peneliti. Setelah proses penghapusan data duplikat, diperoleh sebanyak 1.394 data yang digunakan dalam proses klasifikasi sentimen.

Tabel 3.8 Hasil Evaluasi Model pada Skenario Ground Truth

Algoritma	*Accuracy*	*Precision (Macro Avg)*	*Recall (Macro Avg)*	*F1-Score (Macro Avg)*
Naïve Bayes	56,89%	0,39	0,43	0,40
SVM	57,04%	0,59	0,52	0,54
Random Forest	57,25%	0,61	0,47	0,47

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Berdasarkan Tabel 4., *Random Forest* memperoleh nilai akurasi tertinggi sebesar 57,25%, sedangkan SVM memperoleh nilai F1-Score tertinggi sebesar 0,54. Perbedaan performa antar algoritma relatif kecil karena jumlah data pada skenario *Ground Truth* masih terbatas sehingga model belum mampu mempelajari pola sentimen secara optimal.

3.6.2. Hasil Pengujian Skenario IndoBERT

Skenario IndoBERT menggunakan *dataset* hasil pseudo *labeling* dengan *confidence score* ≥ 0,90. Setelah proses penghapusan data duplikat, diperoleh sebanyak 8.614 data yang digunakan dalam proses klasifikasi sentimen.

Tabel 3.9 Hasil Evaluasi Model pada Skenario IndoBERT

Algoritma	*Accuracy*	*Precision (Macro Avg)*	*Recall (Macro Avg)*	*F1-Score (Macro Avg)*
Naïve Bayes	78,51%	0,54	0,54	0,53
SVM	82,19%	0,79	0,72	0,75
Random Forest	80,88%	0,82	0,64	0,69

Berdasarkan Tabel 4.14, algoritma SVM performa tertinggi dengan ketepatan sebesar 82,19% dan F1-Score sebesar 0,75. Hasil ini menunjukkan bahwa penambahan data berlabel melalui pseudo *labeling* IndoBERT mampu mengoptimalkan kemampuan model dalam mengidentifikasi pola data sentimen dibandingkan skenario *Ground Truth.*

3.6.3. Hasil Pengujian Skenario *InSet Lexicon*

Skenario *InSet Lexicon* menggunakan metode pelabelan berbasis kamus sentimen Bahasa Indonesia. Setelah proses penghapusan data duplikat, diperoleh sebanyak 10.936 data yang digunakan dalam proses klasifikasi.

Tabel 3.10 Hasil Evaluasi Model pada Skenario InSet Lexicon

Algoritma	*Accuracy*	*Precision (Macro Avg)*	*Recall (Macro Avg)*	*F1-Score (Macro Avg)*
Naïve Bayes	77,91%	0,61	0,34	0,30
SVM	81,38%	0,61	0,56	0,58
Random Forest	80,36%	0,65	0,44	0,47

Berdasarkan Tabel 4.18, algoritma SVM kembali memperoleh performa tertinggi dengan ketepatan sebesar 81,38% dan F1-Score sebesar 0,58. Meskipun demikian, distribusi kelas yang tidak merata pada *dataset InSet Lexicon* menyebabkan penurunan kemampuan model dalam mengenali kelas netral dan positif.

3.7. Perbandingan Keseluruhan Performa Model

Tabel 3.11 .Perbandingan Akurasi Seluruh Skenario

Skenario	*Naïve Bayes*	*SVM*	*Random Forest*
Ground Truth	56,89%	57,04%	57,25%
IndoBERT	78,51%	82,19%	80,88%
InSet Lexicon	77,91%	81,38%	80,36%

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Berdasarkan hasil pengujian pada seluruh skenario, algoritma Support Vector Machine (SVM) menunjukkan performa paling stabil dan konsisten. Meskipun Random Forest memperoleh akurasi tertinggi pada skenario Ground Truth sebesar 57,25%, nilainya relatif berdekatan dengan SVM sebesar 57,04%. Pada skenario IndoBERT dan InSet Lexicon, SVM secara konsisten menghasilkan performa terbaik.

Pada skenario IndoBERT, SVM memperoleh akurasi sebesar 82,19% dan Macro F1-Score sebesar 0,75. Hasil ini menunjukkan bahwa penambahan data berlabel melalui pseudo labeling berbasis IndoBERT berpotensi meningkatkan kemampuan klasifikasi sentimen dibandingkan skenario lainnya. Dengan demikian, kombinasi dataset IndoBERT dan SVM menjadi pendekatan paling efektif dalam penelitian ini.

Sebaliknya, Naïve Bayes belum mampu mengenali sentimen positif pada skenario Ground Truth dan IndoBERT, ditunjukkan oleh nilai Precision, Recall, dan F1-Score sebesar 0,00 pada kelas tersebut. Kondisi ini diduga berkaitan dengan ketidakseimbangan distribusi kelas yang menyebabkan model cenderung memprediksi kelas mayoritas. Selain itu, representasi TF-IDF memiliki keterbatasan dalam menangkap hubungan semantik dan konteks kalimat secara menyeluruh.

Pada skenario InSet Lexicon, dominasi sentimen negatif juga menunjukkan adanya ketimpangan distribusi kelas. Pelabelan berbasis leksikon yang mengandalkan pencocokan kata dan skor polaritas diduga kurang mampu memahami konteks kalimat yang kompleks. Hal ini terlihat pada Naïve Bayes yang memperoleh akurasi 77,91%, tetapi hanya menghasilkan Macro F1-Score sebesar 0,30. Perbedaan tersebut menunjukkan bahwa akurasi tinggi belum tentu mencerminkan kemampuan klasifikasi yang merata pada seluruh kelas.

Oleh karena itu, Macro F1-Score menjadi metrik penting karena memberikan bobot yang sama pada setiap kelas sentimen, termasuk kelas minoritas. Metrik ini memberikan gambaran yang lebih menyeluruh mengenai kemampuan model dibandingkan penggunaan akurasi saja.

Selain ketidakseimbangan kelas, kesalahan klasifikasi juga dapat dipengaruhi oleh sarkasme dalam komentar berbahasa Indonesia. Kata seperti “bagus”, “setuju”, atau “hebat” dapat digunakan untuk menyampaikan sindiran atau kritik terhadap kebijakan UU TNI. Karena ketiga algoritma menggunakan fitur TF-IDF yang berfokus pada bobot kata, konteks pragmatik dan maksud tersirat dalam komentar berpotensi tidak teridentifikasi dengan tepat. Keterbatasan ini perlu diperhatikan dalam penelitian selanjutnya, misalnya melalui penggunaan representasi teks yang mampu menangkap konteks bahasa secara lebih mendalam.

3.8. Analisis Sentimen Publik

Berdasarkan hasil pelabelan pseudo labeling IndoBERT, distribusi sentimen komentar YouTube terhadap UU TNI terdiri atas 3.430 komentar negatif (39,82%), 4.770 komentar netral (55,38%), dan 414 komentar positif (4,80%). Dominasi sentimen netral menunjukkan bahwa sebagian besar komentar tidak secara tegas menyatakan dukungan atau penolakan. Sementara itu, proporsi sentimen negatif mencerminkan adanya kritik dan kekhawatiran masyarakat terhadap revisi UU TNI, sedangkan sentimen positif memiliki proporsi paling rendah.



Gambar 3.3 Visualisasi WordCloud

Berdasarkan visualisasi WordCloud pada Gambar 3.3 , kata “rakyat”, “negara”, “DPR”, dan “Indonesia” memiliki frekuensi kemunculan tinggi dalam komentar pengguna YouTube. Hal ini menunjukkan bahwa diskusi banyak berkaitan dengan kebijakan pemerintah, proses legislasi, serta dampak revisi UU TNI terhadap kehidupan berbangsa dan bernegara.

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Kemunculan kata “korupsi”, “koruptor”, “hukum”, dan “jabatan” juga menunjukkan perhatian terhadap penegakan hukum dan tata kelola pemerintahan. Sementara itu, kata “demo”, “mahasiswa”, dan “reformasi” mencerminkan keterkaitan pembahasan dengan dinamika politik dan respons masyarakat terhadap kebijakan tersebut.

Secara keseluruhan, visualisasi WordCloud mendukung hasil analisis sentimen dengan memperlihatkan fokus diskusi pada isu pemerintahan, hukum, dan politik. Dominasi sentimen netral serta tingginya proporsi sentimen negatif menunjukkan bahwa pembahasan revisi UU TNI di YouTube mencakup diskusi yang beragam, termasuk kritik dan kekhawatiran masyarakat. Namun, frekuensi kata dalam WordCloud tidak secara langsung menunjukkan polaritas sentimen suatu komentar.

4. Kesimpulan

Hasil penelitian dan pengujian menunjukkan bahwa analisis sentimen komentar YouTube terkait Undang-Undang Tentara Nasional Indonesia (UU TNI) menggunakan algoritma diperoleh kesimpulan bahwa: 1). Perbandingan ketiga algoritma pada seluruh skenario pelabelan menegaskan bahwa performa klasifikasi sentimen tidak semata-mata ditentukan oleh pilihan algoritma, melainkan sangat dipengaruhi oleh kualitas dan volume data berlabel yang tersedia. *Support Vector Machine (SVM)* terbukti sebagai algoritma yang paling stabil dan unggul secara konsisten di setiap skenario, menunjukkan bahwa kemampuannya membentuk *hyperplane* pemisah kelas lebih tahan terhadap variasi karakteristik data dibandingkan *Naïve Bayes* yang sangat rentan terhadap ketimpangan kelas, maupun *Random Forest* yang cenderung stabil namun kurang sensitif terhadap kelas minoritas. Meningkatnya performa seluruh algoritma pada skenario dengan data lebih besar *(IndoBERT dan InSet Lexicon*) dibandingkan skenario *Ground Truth* membuktikan bahwa keterbatasan data anotasi manual menjadi kendala utama dalam membangun model klasifikasi sentimen yang andal untuk teks berbahasa Indonesia. Dengan demikian, penelitian ini menyimpulkan bahwa pemilihan algoritma yang tepat perlu diiringi strategi perluasan data berlabel yang memadai, dan bahwa evaluasi performa pada data yang tidak seimbang harus mengutamakan metrik yang berimbang antar kelas seperti *F1-Score Macro Average*, bukan akurasi semata, agar kesimpulan mengenai keunggulan suatu model tidak menyesatkan. 2). Konsistennya dominasi sentimen netral dan negatif pada ketiga skenario pelabelan mengindikasikan bahwa Undang-Undang Tentara Nasional Indonesia (UU TNI) belum sepenuhnya mendapatkan kepercayaan dan dukungan terbuka dari publik di ruang digital. Kecilnya proporsi sentimen positif pada seluruh skenario menunjukkan bahwa penerimaan masyarakat terhadap kebijakan ini masih rendah, sementara besarnya porsi sentimen netral mencerminkan kecenderungan publik untuk bersikap hati-hati dan lebih banyak mempertanyakan atau mendiskusikan implikasi kebijakan tersebut ketimbang menyatakan sikap tegas. Temuan ini menegaskan bahwa isu perluasan peran prajurit TNI pada jabatan sipil masih menjadi titik kritis yang memicu kekhawatiran mengenai relasi sipil-militer, sehingga opini publik yang terekam dalam penelitian ini dapat dimaknai sebagai bentuk kontrol sosial masyarakat terhadap kebijakan strategis pemerintah, sekaligus indikasi perlunya komunikasi kebijakan yang lebih transparan dari pemerintah kepada publik.

Reference

1. Ahmad Dzulhijjah, D., Sanjaya, H., Said Wahyudi Hidayat, A., Yulistia Alwanda, A., & Utami, E. (2023). Perbandingan Metode Random Forest dan KNN pada Analisis Sentimen Twitter. *Smart Comp: Jurnalnya Orang Pintar Komputer*, *12*(3). https://doi.org/10.30591/smartcomp.v12i3.5106
2. Alexander Radja Bria, N., & Witanti, A. (2024). ANALISIS SENTIMEN MASYARAKAT INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE TENTANG PILPRES 2024. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(6). https://doi.org/10.36040/jati.v7i6.8184
3. Amrullah, A. (2025). Analisis Sentimen Publik terkait Migrasi Tenaga Kerja Indonesia di Platform X menggunakan SVM-IndoBERT. *Jurnal UJMC*, *11*(1), 53–63.
4. Fathonah, F., & Herliana, A. (2021). Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naïve Bayes. *Jurnal Sains dan Informatika*, *7*(2). https://doi.org/10.34128/jsi.v7i2.331
5. Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *SMATIKA JURNAL*, *10*(02). https://doi.org/10.32664/smatika.v10i02.455
6. google. (2023). *Google Colaboratory*. https://colab.research.google.com/.
7. Haikal, M., Martanto, M., & Hayati, U. (2024). ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI GAME ONLINE PUBG MOBILE MENGGUNAKAN ALGORITMA NAIVE BAYES. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(6). https://doi.org/10.36040/jati.v7i6.8174
8. Hormansyah, D. S., Rahutomo, F., & Aulia, I. (2018). PENERAPAN METODE TF-IDF DAN N-GRAM PADA PENGEMBANGAN APLIKASI CHATBOT BERBASIS LINE UNTUK LAYANAN PUBLIK KESEHATAN DI KOTA MALANG. *Jurnal Informatika Polinema*, *5*(1). https://doi.org/10.33795/jip.v5i1.233
9. Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM. *Jurnal Data Mining dan Sistem Informasi*, *2*(1). https://doi.org/10.33365/jdmsi.v2i1.1021

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

10. Janati Firda. (2025, April 24). *Akademikus Bicara Bahaya TNI Masuk Kampus | tempo.co*. https://www.tempo.co/politik/akademikus-bicara-bahaya-tni-masuk-kampus-1227910
11. Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, *7*(1). https://doi.org/10.24014/coreit.v7i1.12945
12. Putra, A. P., Alivia, A. N., Meilani, M., Azzahra, N. J., & Rakhmawati, N. A. (2023). ANALISIS SENTIMEN WARGA TWITTER TERHADAP GAME SHOPEE COCOKI DENGAN METODE NAIVE BAYES CLASSIFIER. *Jurnal Ilmiah Informatika Komputer*, *28*(2). https://doi.org/10.35760/ik.2023.v28i2.9494
13. Suharsono, J. P., & Nurahman, D. (2024). Pemanfaatan Youtube Sebagai Media Peningkatan Pelayanan Dan Informasi. *Ganaya : Jurnal Ilmu Sosial dan Humaniora*, *7*(1). https://doi.org/10.37329/ganaya.v7i1.3157
14. Suryani, I., Handar, M., Ekasuci, R., Bina Sarana Informatika, U., & Studi Penyiaran, P. (2020). PEMANFAATAN MEDIA SOSIAL SEBAGAI MEDIA PROMOSI BAGI RADIO MERSI 93.9 FM. *Journal Komunikasi*, *11*(1). https://doi.org/10.31294/jkom
15. Tinaliah, T., & Elizabeth, T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, *9*(4). https://doi.org/10.35957/jatisi.v9i4.3586

DOI: https://doi.org/10.69693/ijmst.v5i1.14554
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)