

ANALISIS SENTIMEN PUBLIC TWITTER TERHADAP KEBIJAKAN PEMERINTAH MENGGUNAKAN METODE SVM (STUDI KASUS : RUU TNI)

Fathoni, Ali Ibrahim, Fadia Rizka Mumtaz, Muhammad Azmi Zaky,
M. Jodi Pratama, Iqbal Akbar Kurniawan

Sistem Informasi, Universitas Sriwijaya
Jl. Raya Palembang - Prabumulih No.KM. 32, Indralaya Indah, Kec. Indralaya,
Kabupaten Ogan Ilir, Sumatera Selatan, Indonesia
fathoni@unsri.ac.id

ABSTRAK

Pada 20 Maret 2025, DPR RI mengesahkan RUU TNI menjadi Undang-Undang, yang memicu beragam reaksi publik. Beberapa ketentuan, seperti Pasal 7 ayat (2) huruf b dan Pasal 47 ayat (2), dinilai berpotensi menghidupkan kembali dwifungsi militer dan mengancam prinsip-prinsip demokrasi serta supremasi sipil. Media sosial X (sebelumnya Twitter) menjadi wadah utama masyarakat dalam menyampaikan opini terhadap kebijakan ini. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap UU TNI menggunakan metode *Support Vector Machine* (SVM). Sebanyak 1001 cuitan dikumpulkan melalui teknik *web scraping*, kemudian diproses melalui tahapan *preprocessing* seperti *cleansing*, *tokenisasi*, dan *stemming*. Model SVM mengklasifikasikan sentimen ke dalam kategori positif, netral, dan negatif, dengan akurasi mencapai 94,66%. Hasil evaluasi menunjukkan *recall* tertinggi pada sentimen netral (1,00), diikuti oleh positif (0,98) dan negatif (0,88). Temuan ini mengindikasikan dominasi sentimen negatif dan netral dalam respon publik, sekaligus menjadi masukan penting bagi pemerintah dalam merumuskan kebijakan yang lebih partisipatif dan demokratis.

Kata kunci : Analisis Sentimen; Support Vector Machine; RUU TNI; Media Sosial; Kebijakan Publik

1. PENDAHULUAN

Rancangan Revisi Undang-Undang Tentara Nasional Indonesia (RUU TNI) saat ini sedang menjadi perbincangan hangat di berbagai kalangan masyarakat Indonesia. Hal tersebut disebabkan karena Dewan Perwakilan Rakyat (DPR) RI telah mengesahkan RUU TNI menjadi Undang-Undang (UU) dalam sidang paripurna DPR RI pada 20 Maret 2025. Adapun RUU TNI menuai banyak pro dan kontra disebabkan karena isi pembahasan dari RUU TNI tersebut dianggap berpotensi menghidupkan kembali dwifungsi ABRI dan mengancam kualitas demokrasi. Dalam RUU TNI tersebut juga menimbulkan kekhawatiran dapat membuka jalan bagi prajurit aktif untuk menduduki jabatan sipil [1].

Adapun pasal-pasal problematik yang banyak menuai kontra terdapat pada pasal 7 ayat (2) huruf b secara tersirat memberi legitimasi bahwa TNI merupakan alat negara yang tidak hanya berfungsi di bidang pertahanan, tetapi juga di bidang keamanan negara, dan terdapat pada pasal 47 ayat (2) yang berbunyi “Prajurit dapat menduduki jabatan pada kementerian/lembaga yang membidangi koordinator bidang politik dan keamanan negara, pertahanan negara termasuk dewan pertahanan nasional, kesekretariatan negara yang menangani urusan kesekretariatan presiden dan kesekretariatan militer presiden, intelijen negara, siber dan/atau sandi negara, lembaga ketahanan nasional, pencarian dan pertolongan, narkotika nasional, pengelola perbatasan, penanggulangan bencana, penanggulangan terorisme, keamanan laut, Kejaksaan Republik Indonesia, dan Mahkamah Agung.” [2].

Dari pasal-pasal problematik itulah yang menimbulkan kekhawatiran masyarakat, karena semakin membuka peluang terancamnya supremasi sipil, demokrasi, serta berpotensi terjadinya penyalahgunaan wewenang yang akan berdampak langsung pada Hak Asasi Manusia, dikarenakan tidak adanya regulasi yang ketat dan mekanisme kontrol yang jelas. Sehingga kebijakan ini dapat membawa Indonesia kembali pada pola lama di mana batas antara militer dan sipil menjadi kabur [3].

Hal tersebut akhirnya memicu diskusi publik, terutama pada media sosial Twitter atau X, yang saat ini menjadi wadah populer untuk berbagi opini dan informasi. Penelitian ini memanfaatkan aplikasi X sebagai sumber data untuk menganalisis sentimen masyarakat, dengan mengolah 1000 data cuitan yang ada pada aplikasi X. Analisis sentimen, atau opinion mining, dipilih sebagai metode penelitian. Analisis sentimen merupakan cabang dari pemrosesan bahasa alami (*natural language processing*) yang berfokus pada pengenalan dan klasifikasi emosi serta sikap yang disampaikan melalui bahasa tulis. Tujuannya adalah untuk mendeteksi, memahami, dan mengkategorikan opini atau emosi dalam data teks, baik dalam bentuk positif, negatif, maupun netral. Bidang ini memiliki peran penting dalam mengekstrak opini publik terhadap berbagai entitas dari media sosial dan platform daring lainnya [4].

Seiring dengan pesatnya perkembangan teknologi, penyebaran informasi melalui media sosial berlangsung sangat cepat, memunculkan beragam respons dari masyarakat, baik yang bersifat positif maupun negatif. Untuk menganalisis opini publik yang

umumnya belum terstruktur, diperlukan pemrosesan teks guna memahami sudut pandang masyarakat. Analisis ini penting sebagai bahan pertimbangan bagi pemerintah dalam merumuskan kebijakan yang selaras dengan kondisi dan aspirasi publik [5]. Berdasarkan data tahun 2024, jumlah pengguna internet di Indonesia diperkirakan mencapai 221.563.479 jiwa dari total populasi sebesar 278.696.200 jiwa pada tahun 2023. Hasil Survei Penetrasi Internet Indonesia 2024 mencatat bahwa tingkat penetrasi internet mencapai 79,5%, meningkat sebesar 1,4% dibandingkan dengan tahun sebelumnya [6].

Berbagai studi telah memanfaatkan metode Support Vector Machine (SVM) dalam analisis sentimen terkait kebijakan publik. Salah satunya ditunjukkan pada penelitian dalam [7], yang berhasil mengklasifikasikan opini masyarakat dengan akurasi masing-masing sebesar 82% dan 63%. Sementara itu, penelitian dalam [8] menunjukkan bahwa mengaplikasikan algoritma SVM dalam menganalisis sentimen terhadap kebijakan vaksinasi COVID-19 serta pembatasan sosial, dan berhasil memperoleh nilai presisi dan recall hingga 90%. Di sisi lain, studi dalam [9] menunjukkan bahwa penerapan teknik grid search untuk optimasi parameter pada SVM mampu meningkatkan akurasi analisis sentimen hingga 91%.

Merujuk pada sejumlah penelitian terdahulu, studi ini bertujuan untuk melakukan analisis sentimen terhadap opini publik mengenai Revisi Undang-Undang Tentara Nasional Indonesia (RUU TNI) dengan memanfaatkan algoritma Support Vector Machine (SVM). Algoritma SVM dikenal memiliki performa tinggi serta tingkat akurasi yang unggul dibandingkan metode klasifikasi lainnya karena kemampuannya dalam mengenali pola dan menangani data berdimensi tinggi, yang berkontribusi terhadap hasil klasifikasi yang lebih optimal [10].

Penelitian ini secara khusus mengevaluasi sentimen pengguna platform X terhadap kebijakan revisi RUU TNI menggunakan metode SVM. Struktur penelitian ini mencakup bagian pendahuluan yang mengulas konteks kebijakan dan potensi dampaknya, metodologi yang mencakup tahap pengumpulan data, *praprocessing*, penerapan algoritma SVM, hingga evaluasi performa model. Pada bagian hasil dan pembahasan, dijabarkan metrik performa seperti akurasi, presisi, dan recall, serta analisis tren sentimen yang muncul. Penelitian-penelitian sebelumnya telah membuktikan efektivitas SVM dalam analisis opini publik, seperti yang ditunjukkan pada studi [11] dalam menilai sentimen terhadap kebijakan PPKM, serta dalam studi [12] mengenai kebijakan vaksinasi COVID-19. Keduanya menunjukkan bahwa SVM mampu mengklasifikasikan sentimen secara akurat. Melalui pemahaman terhadap pola sentimen publik, hasil penelitian ini diharapkan dapat memberikan kontribusi sebagai masukan bagi pemerintah dalam mengevaluasi kebijakan dan merumuskan strategi yang lebih responsif terhadap aspirasi masyarakat.

2. TINJAUAN PUSTAKA

Penelitian mengenai analisis sentimen terhadap kebijakan RUU TNI merupakan topik yang relatif baru, mengingat kebijakan ini baru diumumkan dalam waktu dekat. Namun demikian, analisis sentimen terhadap kebijakan pemerintah sebelumnya telah banyak dilakukan menggunakan berbagai metode, salah satunya adalah *Support Vector Machine* (SVM). Dalam penelitian ini, metode SVM dipilih karena telah terbukti efektif dalam mengklasifikasikan. Sentimen pada data teks. Studi dalam [13] menunjukkan bahwa SVM dengan Normalized Poly Kernel mampu memberikan performa yang baik dalam menganalisis sentimen publik terhadap kebijakan pemerintah terkait penanganan COVID-19. Data yang digunakan berasal dari Twitter, yang menjadi salah satu platform utama opini publik.

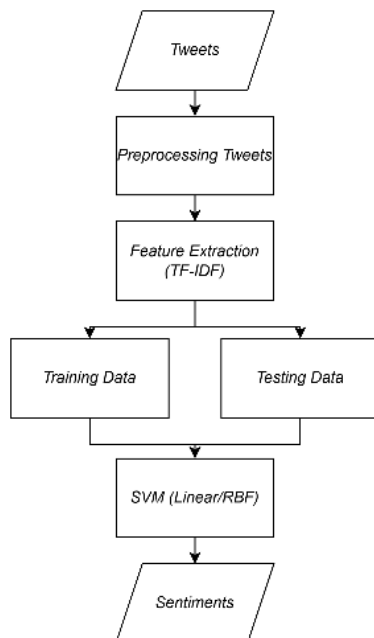
Penelitian dalam [14] menunjukkan bahwa penggunaan SVM untuk melakukan analisis sentimen terhadap kebijakan publik seperti Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) mencapai rata-rata akurasi sebesar 77,2%, presisi sebesar 83,3%, recall sebesar 68,7%, dan F-measure sebesar 75,11%. Studi ini menegaskan bahwa SVM dapat digunakan untuk menganalisis opini publik terhadap kebijakan pemerintah berdasarkan data Twitter, meskipun tidak dibandingkan secara langsung dengan metode lain seperti Naive Bayes. Sejalan dengan hal tersebut, penelitian dalam [15] menunjukkan bahwa pada analisis sentimen menteri-menteri Indonesia di Twitter menunjukkan metode SVM memiliki akurasi yang lebih tinggi dibandingkan Naive Bayes dalam mengklasifikasikan sentimen pengguna media sosial, dengan akurasi masing-masing sebesar 89,60% dan 85,74%. Hasil ini menunjukkan bahwa SVM lebih efektif dalam menangkap opini publik dalam konteks ini, meskipun penerapannya pada kebijakan nasional yang lebih luas memerlukan penelitian lebih lanjut.

Belakangan ini pemberitaan di media sosial juga menunjukkan kegelisahan masyarakat terhadap hasil berbagai kebijakan pemerintah yang menunjukkan hasil sentimen negatif mendominasi. Laporan dalam [16] menunjukkan bahwa sebagian besar masyarakat mempunyai pandangan negatif terhadap kebijakan Pemerintah, dan banyak komentar yang menyatakan ketidakpuasan. Sementara itu, Syamsu Rizal, anggota Komite I dari Fraksi PKB, juga menegaskan, proses pembahasan RUU TNI harus lebih transparan agar tidak menimbulkan kontroversi yang lebih besar. Oleh karena itu, penggunaan *Support Vector Machine* untuk analisis sentimen dapat membantu pengambil kebijakan guna mendapatkan pemahaman yang lebih mendalam mengenai dinamika opini publik dan pada akhirnya menjadi pedoman pemerintah dalam merumuskan kebijakan yang lebih sesuai dengan keinginan masyarakat [17].

Dengan mengacu pada berbagai penelitian yang ada, dapat disimpulkan bahwa SVM adalah metode yang tepat untuk analisis sentimen terhadap kebijakan

publik yang kontroversial seperti RUU TNI. Keunggulan SVM dalam mengolah data besar dan kompleks menjadikannya pilihan utama dalam penelitian ini untuk memberikan gambaran yang lebih akurat mengenai persepsi masyarakat terhadap kebijakan tersebut.

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang umum digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi. Algoritma ini berfungsi dengan menemukan hyperplane terbaik yang dapat memisahkan data ke dalam dua atau lebih kelas, sambil memaksimalkan jarak atau margin antara kelas-kelas tersebut [18]. Konsep dasar SVM melibatkan penggunaan fungsi kernel untuk mentransformasikan data ke dalam ruang berdimensi lebih tinggi sehingga memungkinkan pemisahan yang lebih efektif antar kelas [7]. Ilustrasi mengenai tahapan kerja SVM dapat dilihat pada Gambar 1.



Gambar 1. Proses Kerja SVM

Dalam konteks analisis sentimen, SVM diawali dengan proses pengumpulan data dari media sosial seperti Twitter, yang diperoleh melalui teknik crawling untuk mendapatkan data mentah yang akan dianalisis [19]. Setelah data terkumpul, dilakukan proses *preprocessing* yang mencakup pembersihan teks seperti menghapus *stopwords*, melakukan *stemming*, *lemmatization*, dan *tokenisasi* guna menghasilkan data yang terstruktur dan siap diolah. Data yang telah diproses kemudian diubah menjadi bentuk numerik menggunakan teknik seperti TF-IDF atau Word Embeddings, yang berguna untuk memberikan bobot yang sesuai terhadap kata-kata penting dalam data sehingga meningkatkan kinerja model SVM [20].

Langkah selanjutnya adalah membagi data menjadi dua bagian, yaitu data latih (training set) dan data uji (testing set), agar model yang dibangun

mampu melakukan generalisasi dengan baik pada data yang belum pernah dilihat sebelumnya [21]. Pelatihan model dilakukan dengan memilih jenis kernel yang sesuai, seperti Linear, RBF, atau Polynomial, tergantung pada karakteristik data yang dianalisis [22]. Evaluasi model dilakukan menggunakan metrik-metrik seperti akurasi, precision, recall, dan F1-score untuk mengukur tingkat keberhasilan klasifikasi sentimen oleh model SVM [23]. Setelah memperoleh model dengan performa terbaik, model tersebut dapat diterapkan pada data baru untuk melakukan analisis sentimen secara otomatis. Hasil analisis ini kemudian dapat dimanfaatkan sebagai bahan pertimbangan dalam proses pengambilan keputusan, khususnya dalam ranah kebijakan publik [24].

3. METODE PENELITIAN

Metode penelitian adalah sebuah prosedur sistematis yang disusun untuk melaksanakan penelitian guna mencapai tujuan yang diinginkan. Adapun alur rancangan penelitian yang dilakukan dapat dilihat pada Gambar 2.



Gambar 2. Alur penelitian

3.1. Data Selection

Pada tahap ini, data dikumpulkan menggunakan teknik crawling, yaitu metode pengambilan data dari sebuah situs web dengan memasukkan *Uniform Resource Locator (URL)* sebagai titik awal. URL ini digunakan untuk menelusuri seluruh *hyperlink* yang terdapat dalam situs tersebut. Setelah itu, dilakukan proses pengindeksan untuk menemukan kata-kata dalam dokumen pada setiap tautan yang ditemukan. Data yang dikumpulkan berupa *tweet* dari media sosial Twitter, yang diperoleh melalui pemanfaatan Twitter API. Untuk dapat mengakses *tweet* tersebut, diperlukan izin akses berupa *consumer key*, *consumer secret*, *access token*, dan *access token secret* [12]. Kata kunci yang menjadi keyword adalah RUU TNI. *Crawling* data diambil dari data dengan *tweets* tanggal 1 Februari 2005 sampai dengan 25 Maret 2025 dengan total perolehan tweets sejumlah 1.001 tweets.

3.2. Data Preprocessing

Setelah mendapatkan 1001 data *tweets* dengan kata kunci RUU TNI. Data akan masuk ke tahap selanjutnya yaitu *preprocessing* data. *Preprocessing* data dilakukan dalam enam tahapan yaitu :

- Cleansing* : tahap ini bertujuan untuk mengurangi gangguan atau noise dalam data tweet yang telah dikumpulkan, seperti menghapus simbol, tautan URL, tanda pagar (#), serta *mention* (@).
- Case Folding* : proses ini mengonversi seluruh huruf kapital menjadi huruf kecil guna mencegah duplikasi atau keragaman bentuk kata yang seharusnya sama.
- Tokenization* : merupakan proses pemisahan teks atau kalimat menjadi potongan-potongan kata berdasarkan spasi dan tanda baca tertentu, seperti tanda kutip tunggal.
- Stopword Removal* : di tahap ini, kata-kata yang tidak memiliki nilai penting dalam analisis, seperti “dan”, “yang”, “ke”, serta kata hubung lainnya akan dihilangkan dari data.
- Normalization* : proses ini mengubah kata-kata tidak baku dalam tweet menjadi bentuk kata yang sesuai dengan kaidah bahasa yang benar.
- Stemming* : pada langkah ini, kata-kata yang memiliki imbuhan akan dikembalikan ke bentuk dasar atau bentuk asalnya, agar mempermudah analisis data lebih lanjut.

3.3. Labelling

Tahap ini bertujuan untuk menetapkan kategori atau kelas dari setiap ulasan yang diperoleh, apakah termasuk dalam sentimen positif, negatif, atau netral.

3.4. Data Transformation

Pada proses ini, label sentimen yang semula berbentuk teks seperti positif, netral, dan negatif diubah menjadi format numerik. Label positif dikodekan sebagai angka 2, netral sebagai 1, dan negatif sebagai 0.

3.5. Text Mining

Tahapan ini melibatkan pemrosesan data teks melalui klasifikasi tweet menggunakan algoritma *Support Vector Machine* (SVM). Tujuannya adalah untuk mengolah data mentah menjadi informasi yang bernilai dan dapat dianalisis lebih lanjut.

3.6. Evaluation

Pada tahap akhir, dilakukan proses evaluasi untuk menilai validitas model yang dihasilkan oleh algoritma yang telah diterapkan. Kinerja algoritma diukur menggunakan tiga indikator utama, yaitu *Accuracy*, *Precision*, dan *F1-Score* [19].

4. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan analisis sentimen dari pengguna aplikasi “X” terhadap kebijakan pemerintah terkait RUU TNI dengan menerapkan algoritma *Support Vector Machine* (SVM). Sentimen diklasifikasikan ke dalam tiga kategori, yaitu positif, netral, dan negatif, menggunakan bahasa pemrograman Python.

4.1. Data Selection

Proses ini dilakukan dengan metode *crawling data* pada platform “X”, dengan menargetkan cuitan yang mengandung kata kunci “RUU TNI”. Dari proses ini, berhasil dihimpun sebanyak 1001 tweet. Tabel 1 di bawah ini menampilkan dataset awal yang diperoleh melalui teknik *crawling* tersebut.

Tabel 1. Dataset awal hasil *crawling data*

No	Date	User	Tweet
1	24-03-2025, 23:49:54	YuvEbig	bukannya Pemerintahan sekarang yang cabut atau melemah atau seenggaknya RUU TNI & POLRI dicabut yang dapet malah ini gimana ga rentan dipermainkan narasinya coba? Semoga aja Oposisi tidak sebodoh yang sudah-sudah
2	24-03-2025, 23:47:02	https_jjisvng	@tropicsouhls oalaa gitu ya kak. sejujurnya smnjak ruu tni di sahkan kita tuh jdi susah buat mikir positif ke aparat' takut mreka macem'atau gimana gitu kan. tpi intinya dmn mn kt jg hrs ttp berhati'sm siapapun itu. anw thanks ya kak infonyaa
3	24-03-2025, 23:49:54	Rosalin020292	@Heraloebs @kamto_adi @prabowo @DPR_RI batalkan itu RUU baru TNI sebelum semuanya terlambat. Jika dibiarkan bukan hanya emosi mahasiswa bisa menyebar di masyarakat... Hingga ke mancanegara ... Hingga negara hancur korban dimana2 diakhir rusia dan Amerika akan mempersenjatai rakyat sipil.
4	24-03-2025, 23:49:54	ImperfectDilan	pemerintah perlu lebih tegas ya makanya tegas dong tolak RUU TNI & POLRI terus urus UU perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol
5	24-03-2025, 23:49:54	grok	@alfatih212426 Rakyat bisa menentang RUU TNI Polri dan Kejaksaan lewat demo damai petisi lobi DPR kampanye media sosial atau gugatan hukum ke MK. RUU ini kontroversial karena diduga memperluas kuasa dan kurangi pengawasan sipil. TNI disahkan 20 Maret 2025 Polri dan Kejaksaan masih https://t.co/DQXZmxo7Me

4.2. Data Preprocessing

Data yang telah dikumpulkan akan diproses melalui tahap preprocessing. Tujuan dari tahap ini adalah untuk membersihkan, mengorganisasikan, dan mempersiapkan data sehingga dapat diproses dengan benar. Proses preprocessing dilakukan menggunakan Google Collab dengan menggunakan bahasa Python.

a. Cleansing

Sebelum proses *cleansing*, data penuh dengan gangguan; namun, setelah proses *cleansing*, data menjadi lebih teratur dan bersih, yang membuatnya lebih mudah untuk diproses pada tahap berikutnya.

Tabel 2. *Cleansing*

Sebelum Cleansing	Setelah Cleansing
pemerintah perlu lebih tegas ya makanya tegas dong tolak RUU TNI & POLRI terus urus UU perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol	pemerintah perlu lebih tegas ya makanya tegas dong tolak RUU TNI & POLRI terus urus UU perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol

b. Case Folding

Pada tahap ini, semua huruf dari setiap kata dalam dataset sebelumnya diubah menjadi huruf lowercase.

Tabel 3. *Case Folding*

Sebelum Case Folding	Setelah Case Folding
pemerintah perlu lebih tegas ya makanya tegas dong tolak RUU TNI & POLRI terus urus UU perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol	pemerintah perlu lebih tegas ya makanya tegas dong tolak ruu tni amp polri terus urus uu perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol

c. Tokenization

Pada bagian ini, teks dibagi menjadi kata kata individu berdasarkan spasi.

Tabel 4. *Tokenization*

Sebelum Tokenization	Setelah Tokenization
pemerintah perlu lebih tegas ya makanya tegas dong tolak ruu tni amp polri terus urus uu perampasaan aset koruptor ikutin suara rakyat kita semua juga ga bakal tuh gaduhin istana kalo kerjanya emang bener mah tolol	['pemerintah', 'perlu', 'lebih', 'tegas', 'ya', 'makanya', 'tegas', 'dong', 'tolak', 'ruu', 'tni', 'amp', 'polri', 'terus', 'urus', 'uu', 'perampasaan', 'aset', 'koruptor', 'ikutin', 'suara', 'rakyat', 'kita', 'semua', 'juga', 'ga', 'bakal', 'tuh', 'gaduhin', 'istana', 'kalo', 'kerjanya', 'emang', 'bener', 'mah', 'tolol']

d. Stopword Removal

Pada tahapan stopwords removal, kata-kata yang sering digunakan namun tidak terlalu penting untuk proses analisis dihilangkan.

Tabel 5. *Stopword Removal*

Sebelum Stopword Removal	Setelah Stopword Removal
['pemerintah', 'perlu', 'lebih', 'tegas', 'ya', 'makanya', 'tegas', 'dong', 'tolak', 'ruu', 'tni', 'amp', 'polri', 'terus', 'urus', 'uu', 'perampasaan', 'aset', 'koruptor', 'ikutin', 'suara', 'rakyat', 'kita', 'semua', 'juga', 'ga', 'bakal', 'tuh', 'gaduhin', 'istana', 'kalo', 'kerjanya', 'emang', 'bener', 'mah', 'tolol']	['pemerintah', 'ya', 'tolak', 'ruu', 'tni', 'polri', 'urus', 'uu', 'perampasaan', 'aset', 'koruptor', 'ikutin', 'suara', 'rakyat', 'gaduhin', 'istana', 'kerjanya', 'tolol']

e. Normalization

Kata tidak baku diubah menjadi baku pada tahap ini.

Tabel 6. *Normalization*

Sebelum Normalization	Setelah Normalization
['pemerintah', 'ya', 'tolak', 'ruu', 'tni', 'polri', 'urus', 'uu', 'perampasaan', 'aset', 'koruptor', 'ikutin', 'suara', 'rakyat', 'gaduhin', 'istana', 'kerjanya', 'tolol']	['pemerintah', 'iya', 'tolak', 'ruu', 'tni', 'amp', 'polri', 'mengurus', 'uu', 'perampasan', 'aset', 'koruptor', 'ikuti', 'suara', 'rakyat', 'tidak', 'itu', 'mengganggu', 'istana', 'kalau', 'pekerjaannya', 'memang', 'benar', 'mah', 'tolol']

f. Stemming

Jika terdapat imbuhan pada kata tersebut, maka akan dikembalikan ke bentuk aslinya.

Tabel 7. *Stemming*

Sebelum Stemming	Sesudah Stemming
['pemerintah', 'iya', 'tolak', 'ruu', 'tni', 'amp', 'polri', 'mengurus', 'uu', 'perampasaan', 'aset', 'koruptor', 'ikuti', 'suara', 'rakyat', 'tidak', 'itu', 'mengganggu', 'istana', 'kalau', 'pekerjaannya', 'memang', 'benar', 'mah', 'tolol']	['perintah', 'iya', 'tolak', 'ruu', 'tni', 'amp', 'polri', 'urus', 'uu', 'rampas', 'aset', 'koruptor', 'ikut', 'suara', 'rakyat', 'tidak', 'itu', 'ganggu', 'istana', 'kalau', 'kerja', 'memang', 'benar', 'mah', 'tolol']

4.3. Data Transformation

Data yang telah melalui proses preprocessing akan masuk ke tahap transformation. Transformasi data dilakukan untuk meningkatkan kualitas data, menyamakan skala, dan memastikan kompatibilitas dengan metode atau model yang digunakan. Tujuan dari transformasi data adalah untuk mengubah format, struktur, atau nilai data sehingga lebih sesuai untuk analisis atau pemrosesan berikutnya.

a. *Labelling*

Peneliti memberikan label pada seluruh data dari tahap preprocessing untuk menentukan ulasan menjadi kelas positif, negatif, atau netral. Tabel di bawah menunjukkan kata yang termasuk kata positif atau negatif. Kata netral tidak disertakan karena akan dikategorikan hanya jika kata tersebut tidak termasuk kata positif atau negatif.

Tabel 8. *Labelling*

<i>Positive word</i>	<i>Negative Word</i>
['bagus', 'baik', 'puas', 'senang', 'mantap', 'suka', 'hebat', 'bijak', 'setuju', 'sah', 'dukung', 'sudah']	['buruk', 'tidak', 'tolak', 'demo', 'benci', 'rusak', 'kecewa', 'kurang', 'salah', 'cabutuutni', 'aksi', 'indonesiagelap', 'rakyat', 'polri']

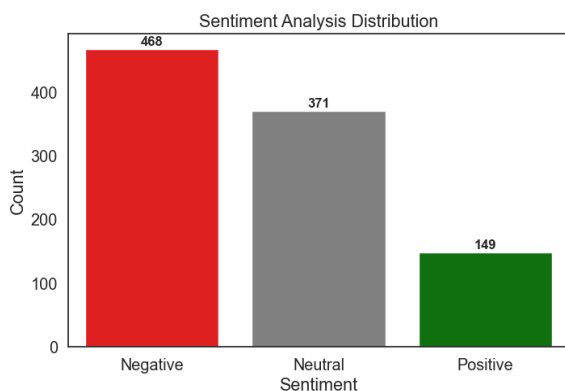
b. *Label Encoding*

Setelah proses labelling selesai, setiap kelas yang diberi label akan diubah menjadi angka agar dapat diproses oleh model machine learning. Tabel di bawah menunjukkan output label encoding dari kelas-kelas yang ada saat ini.

Tabel 9. *Label Encoding*

<i>Label</i>	<i>Label Encoding</i>
<i>Positive</i>	2
<i>Netral</i>	1
<i>Negatif</i>	0

Proses *label encoding* menghasilkan jumlah sentimen negatif sebanyak 468 data, sentimen netral sebanyak 371 data dan sentimen positif sebanyak 149 data. Grafik visualisasi distribusi sentimen dari hasil proses *label encoding* ditunjukkan oleh gambar 3.



Gambar 3. Grafik distribusi sentimen

4.4. *Text Mining*

Pada tahap *text mining*, proses klasifikasi dilakukan dengan menerapkan algoritma SVM terhadap data teks yang telah melalui tahap *preprocessing*. Data yang digunakan dalam klasifikasi terdiri dari fitur yang diekstraksi menggunakan TF-IDF serta fitur tambahan berupa jumlah kata positif dan negatif dalam teks. Sementara itu, label ulasan yang sebelumnya berbentuk kategori (positif, negatif,

dan netral) telah dikonversi ke dalam format numerik untuk dapat digunakan dalam model SVM.

Selanjutnya, dilakukan optimasi hyperparameter menggunakan GridSearchCV untuk menentukan kombinasi parameter terbaik, seperti kernel, nilai C, dan gamma. Model yang telah dilatih dievaluasi menggunakan metrik akurasi, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

	precision	recall	f1-score	support
0	1.00	0.88	0.93	104
1	0.85	1.00	0.92	83
2	1.00	0.98	0.99	94
accuracy			0.95	281
macro avg	0.95	0.95	0.95	281
weighted avg	0.95	0.95	0.95	281

Akurasi Model SVM: 0.9466192170818505

Gambar 4. Hasil program klasifikasi SVM

Gambar di atas menunjukkan hasil dari laporan klasifikasi, skor akurasi, dan tabel *confusion matrix*. Tingkat akurasi model SVM yang digunakan menghasilkan akurasi sebesar 94,66%.

4.5. *Evaluation*

Evaluasi dilakukan untuk mengukur performa model Support Vector Machine (SVM) dalam mengklasifikasikan sentimen publik terhadap kebijakan Revisi Undang-Undang TNI. Evaluasi dilakukan berdasarkan metrik akurasi, *precision*, *recall*, dan *F1-score*, serta analisis terhadap distribusi prediksi dalam *confusion matrix*.

Model SVM menunjukkan akurasi sebesar 94,66%, dengan nilai recall tertinggi pada kelas netral sebesar 1,00, yang berarti semua tweet dengan sentimen netral berhasil diklasifikasikan dengan benar. Sementara itu, recall untuk kelas negatif sebesar 0,98 dan positif sebesar 0,88, menunjukkan bahwa model mampu mengenali hampir semua data sentimen negatif, tetapi sedikit kurang optimal dalam mengenali data sentimen positif.

Precision, *recall*, dan *F1-score* dihitung untuk mengevaluasi keseimbangan antara prediksi positif dan negatif yang benar. *Precision* mengukur tingkat ketepatan prediksi positif yang benar dibandingkan dengan seluruh prediksi positif yang dibuat oleh model:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (1)$$

Sedangkan *recall* menunjukkan kemampuan model dalam mendeteksi seluruh data positif yang sebenarnya ada:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

F1-score digunakan sebagai metrik keseluruhan untuk menyeimbangkan *precision* dan *recall*, dengan rumus:

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Meskipun demikian, distribusi data yang tidak seimbang (dengan dominasi data negatif dan netral dibanding positif) mempengaruhi performa klasifikasi pada kelas minoritas. Nilai *F1-score* rata-rata sebesar 0,95 menunjukkan bahwa model memiliki performa yang stabil dan baik secara keseluruhan, meskipun terdapat indikasi model *overfitting* terhadap kelas netral karena *recall*-nya yang sempurna namun kurang variasi prediksi.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa metode Support Vector Machine (SVM) efektif dalam mengklasifikasikan sentimen publik terhadap RUU TNI berdasarkan data dari Twitter. Dengan akurasi 94,66%, model mampu mengenali sentimen netral secara sempurna 1,00, serta sentimen negatif dan positif dengan *recall* masing-masing 0,98 dan 0,88. Hasil klasifikasi memperlihatkan bahwa respons publik didominasi oleh sentimen netral dan negatif, mencerminkan kekhawatiran masyarakat terhadap isi RUU tersebut.

Meski hasilnya cukup baik, distribusi data yang tidak seimbang menyebabkan model cenderung lebih akurat pada sentimen netral. Oleh karena itu, penelitian lanjutan disarankan untuk menyeimbangkan data latih, mengoptimalkan parameter model, serta mempertimbangkan penggunaan fitur berbasis konteks seperti *word embeddings*. Hasil penelitian ini dapat menjadi masukan bagi pemerintah dalam menyusun kebijakan dengan memberikan ruang bagi masyarakat dalam andil pembentukan kebijakan yang lebih sesuai dengan aspirasi publik.

DAFTAR PUSTAKA

- [1] M. Z. Habibie, Y. S. P. Sihalo, R. M. Andeka, M. A. Kiansantang, C. Irawan, dan K. Amaliah, "YAYASAN LEMBAGA BANTUAN HUKUM INDONESIA (YLBHI) MENOLAK UPAYA MENGHIDUPKAN KEMBALI DWI FUNGSI MELALUI REVISI UU TNI (TENTARA NEGARA INDONESIA)," *Quantum Juris: Jurnal Hukum Modern*, vol. 7, no. 2, 2025.
- [2] U. Hamid, "Menimbang Kadar Dwifungsi (UU) TNI," *Kompas.id*, 24 Maret 2025. [Online]. Tersedia: <https://www.kompas.id/artikel/menimbang-kadar-dwifungsi-uu-tni>
- [3] H. K. Hoetomo and W. Thukul, "Upaya keluar kandang (barak) oleh TNI: Reformasi yang berputar di tempat?," unpublished, Mar. 2025. [Online],doi: <https://doi.org/10.13140/RG.2.2.34172.65927>.
- [4] M. Kumar, L. Khan, and H.-T. Chang, "Evolving techniques in sentiment analysis: a comprehensive review," *PeerJ Computer Science*, vol. 11, p. e2592, Jan. 2025.
- [5] M. S. Hasibuan dan A. Serdano, "Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan Support Vector Machine dan Naïve Bayes," *Jurnal Riset Sains dan Teknologi*, vol. 6, no. 2, pp. 199–204, 2022, doi: <https://doi.org/10.30595/jrst.v6i2.15145>.
- [6] T. Pratama, R. Nugroho, dan M. Aditya, "Comparison of SVM, Random Forest, and KNN in Sentiment Analysis of Online Learning Policy," *International Journal of Educational Data Science*, vol. 19, no. 2, hlm. 120–135, 2021.
- [7] M. Yasir, M. G. Haque, R. Suraji, dan I. Sastrodiharjo, "Analisis Sentimen Terhadap Kontroversi Fatwa MUI Nomor 83 Tahun 2023 Tentang Pemboikotan Produk yang Terafiliasi Israel," *Jurnal Ekonomi Manajemen Sistem Informasi*, vol. 5, hlm. 409–422, 2024.
- [8] Q. A. Chairunnisa, Y. Herdiyeni, M. K. D. Hardhienata, and J. Adisantoso, "Analisis sentimen pengguna Twitter terhadap program vaksinasi COVID-19 di Indonesia menggunakan algoritme Support Vector Machine," *Jurnal Ilmu Komputer dan Agri-Informatika*, vol. 9, no. 1, pp. 79–89, 2022.
- [9] Dendy, A., & Sugihartono, T., "Perbandingan Teknik Optimasi Grid Search dan Randomized Search dalam Meningkatkan Akurasi Metode Klasifikasi SVM," *SKANIKA*, vol. 8, no. 1, pp. 13–22, 2025.
- [10] Ermawan, B., & Cahyono, N., "Optimasi Metode Klasifikasi Menggunakan FastText dan Grid Search," *JIKO*, vol. 9, p. 226, 2025.
- [11] Pradhana, R. M., "Analisis Sentimen Publik terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Skala Mikro Menggunakan Algoritma Support Vector Machine Studi Kasus Twitter," Doctoral dissertation, Universitas Dinamika, 2021.
- [12] Hasibuan, M. S., & Suhardi, T., "Analysis of Covid-19 Vaccine Policy Sentiment Using SVM and C4.5," *Jurnal Teknik Elektro dan Komputer TRIAC*, vol. 9, no. 2, pp. 67–69, 2022.
- [13] L. N. Pradany and C. Fatichah, "Analisa Sentimen Kebijakan Pemerintah pada Konten Twitter Berbahasa Indonesia Menggunakan SVM dan K-Medoid Clustering," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)*, Universitas Brawijaya, vol. 5, no. 10, pp. 3894–3901, 2021.
- [14] D. W. Syahputra, B. Rahayudi, and L. Muflikhah, "Analisis Sentimen Twitter terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat menggunakan Metode Support Vector Machine," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-*

- PTIIK), vol. 6, no. 4, pp. 1760–1767, 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [15] R. Pebrianto, S. N. Nugraha, A. Latif, and M. R. Firdaus, "Analisis Sentimen Twitter terhadap Menteri Indonesia dengan Algoritma Support Vector Machine dan Naive Bayes," *Jurnal Informatika Universitas Pamulang*, vol. 17, pp. 1–12, 2022.
- [16] H. Yaputra and M. R. Muzzaki, "Tagar 'Indonesia Gelap' mendapatkan sentimen negatif," *Tempo.co*, 2023. [Online]. Available: <https://www.tempo.co/politik/tagar-indonesia-gelap-dapat-81-persen-sentimen-negatif-1221552> [Accessed: Apr. 1, 2025].
- [17] Kompas TV, "Syamsu Rizal: Proses Pembahasan RUU TNI Harus Lebih Transparan," Kompas TV, 2023. [Online]. Available: <https://www.kompas.tv/nasional/581077/anggota-komisi-i-dpr-fraksi-pkb-ini-akui-pembahasan-revisi-uu-tni-kurang-transparan>. [Accessed: Apr. 1, 2025].
- [18] M. Rahman and R. Fadhil, "Support Vector Machine for Text Classification: A Comprehensive Review," *Journal of Artificial Intelligence Research*, vol. 45, pp. 112–130, 2021.
- [19] D. Suryana and N. Lestari, "Text Preprocessing Techniques for Sentiment Analysis in Social Media Data," *Journal of Computer Science and Information Systems*, vol. 12, no. 4, pp. 221–238, 2023.
- [20] P. Wijaya and T. Prasetyo, "TF-IDF vs. Word Embeddings: Which One is Better for Sentiment Analysis?," in *Proc. Int. Conf. Data Science*, pp. 33–45, 2022.
- [21] R. Hakim and S. Rahman, "Training and Testing Strategies in Machine Learning-Based Sentiment Analysis," *International Journal of Artificial Intelligence and Data Mining*, vol. 11, no. 2, pp. 87–102, 2024.
- [22] B. Santoso, A. Nugraha, and M. Setyawan, "Comparison of Kernel Types in SVM for Sentiment Analysis," *Journal of Intelligent Systems and Applications*, vol. 9, no. 4, pp. 211–226, 2025.
- [23] M. Firdaus and A. Ramadhan, "Evaluating Sentiment Classification Models: Metrics and Best Practices," *Journal of Data Science and Machine Learning*, vol. 13, no. 1, pp. 45–61, 2023.
- [24] T. Nugroho, F. Mahendra, and D. Satria, "Application of Sentiment Analysis for Public Policy Decision Making," *International Journal of Government and Technology Studies*, vol. 8, no. 3, pp. 190–204, 2025.