Jurnal Ilmiah Wahana Pendidikan, June 2026, 12(6.B), 80-89
DOI: https://jurnal.peneliti.net/index.php/JIWP/article/view/13271
p-ISSN: 2622-8327 e-ISSN: 2089-5364
Accredited by Directorate General of Strengthening for Research and Development
Available online at https://jurnal.peneliti.net/index.php/JIWP

Perbandingan *Support Vector Machine* dan *Random Forest* dalam Analisis Sentimen Komentar YouTube Terkait Isu Hak Veto Amerika Serikat

Raival Maulidan Muhamad Akbar[1] , Pandapotan Kristian Sitorus[2], Fergiano Deren Ryandi[3], Muhammad Rizqi Warsita[4] , Chaerur Rozikin[5]

[1,2,3,4,5]Universitas Singaperbangsa Karawang

Received: 20 Mei 2026
Revised: 29 Mei 2026
Accepted: 11 Juni 2026

Abstract

This study aims to compare the performance of two classification algorithms Random Forest and Support Vector Machine (SVM) with a sigmoid kernel in conducting sentiment analysis on YouTube comments related to the issue of the United States' veto power. The dataset consists of 3,363 comments that have undergone comprehensive preprocessing steps (cleaning, normalization, tokenization, etc.) and were manually labeled into two sentiment classes: positive and negative. The findings indicate that SVM provides a more balanced classification across both sentiment categories, although its overall accuracy is slightly lower at 88.00%. In contrast, Random Forest achieves the highest accuracy at 89.33%, making it superior in terms of overall predictive performance. Therefore, SVM is more suitable when balanced class performance is the priority, whereas Random Forest is preferable when maximizing classification accuracy is the primary objective.

Keywords: *Sentiment Analysis, Support Vector Machine, Random Forest; YouTube Comments, United States Veto Power.*

(*) Corresponding Author: raivalm@gmail.com[1] , pandachristian27@gmail.com[2], fergianoderenryandi@gmail.com[3], rizqyyahoo1@gmail.com[4] , chaerur.rozikin@staff.unsika.ac.id[5]

How to Cite: Muhamad Akbar, R. M., Sitorus, P. K., Ryandi, F. D., Warsita, M. R., & Rozikin, C. (2026). Perbandingan Support Vector Machine dan Random Forest dalam Analisis Sentimen Komentar YouTube Terkait Isu Hak Veto Amerika Serikat. *Jurnal Ilmiah Wahana Pendidikan*, *12*(6.B), 80-89. Retrieved from https://jurnal.peneliti.net/index.php/JIWP/article/view/13271.

PENDAHULUAN

Penelitian ini memberikan kontribusi penting dalam ranah akademis maupun praktis dengan menghadirkan analisis komparatif antara algoritma Support Vector Machine (SVM) dan Random Forest dalam mengklasifikasikan sentimen publik terkait isu politik global, khususnya hak veto Amerika Serikat di PBB. Di tengah lebih dari 167 juta pengguna media sosial di Indonesia pada tahun 2024 sekitar 60% dari populasi YouTube telah berkembang menjadi ruang diskusi yang memengaruhi opini publik, sehingga respons masyarakat terhadap isu veto AS menjadi relevan untuk ditelaah secara ilmiah. Dengan memanfaatkan 3.363 komentar YouTube yang telah melalui rangkaian preprocessing (pembersihan, normalisasi, tokenisasi, dan lainnya), penelitian ini mengelompokkan komentar ke dalam kelas sentimen positif, negatif, dan netral.

Evaluasi model dilakukan dengan metrik akurasi, presisi, recall, dan F1-score untuk memastikan penilaian yang objektif. Hasilnya menunjukkan bahwa Random Forest meraih akurasi tertinggi, sementara SVM tampil lebih stabil dan seimbang di kedua kelas sentimen. Secara akademis, penelitian ini memperluas literatur analisis sentimen pada media sosial dan mengisi kekosongan studi terkait

isu politik global melalui perbandingan langsung SVM dan Random Forest. Dari sisi praktis, temuan ini membantu memahami persepsi publik terhadap kebijakan internasional dan dapat menjadi rujukan untuk analisis opini pada isu global lainnya.

TINJAUAN PUSTAKA

1) Research Gap

Penelitian sebelumnya banyak berfokus pada domain komersial dan isu politik domestik, sehingga kajian mengenai hak veto AS di PBB masih minim dan umumnya hanya menguji satu algoritma. Kesenjangan ini menunjukkan perlunya analisis sentimen pada isu politik internasional serta perbandingan langsung antara SVM dan Random Forest. Penelitian ini mengisi kekosongan tersebut dengan menganalisis sentimen komentar YouTube tentang hak veto AS dan membandingkan kinerja kedua algoritma. Hasilnya memberikan gambaran jelas tentang kekuatan masing-masing metode serta dasar empiris untuk memetakan opini publik dan memilih algoritma yang tepat.

2) State of the Art

Kajian terkini menunjukkan bahwa *Support Vector Machine* (SVM) masih menjadi salah satu algoritma unggulan dalam analisis teks berdimensi tinggi karena mampu menjaga margin antar kelas, sedangkan *Random Forest* (RF) semakin banyak digunakan untuk mengurangi risiko *overfitting* melalui pendekatan *ensemble learning*. Beberapa penelitian terbaru (misalnya pada analisis ulasan aplikasi digital dan komentar media sosial) menunjukkan bahwa RF seringkali unggul dalam hal akurasi global, sementara SVM memberikan hasil yang lebih seimbang antar kelas. Penelitian ini berkontribusi pada *state of the art* dengan membandingkan kedua algoritma pada konteks isu politik global yang belum banyak dikaji, sehingga memberikan wawasan baru tentang keunggulan relatif masing-masing algoritma dalam domain opini publik internasional.

3) Analisis Sentimen

Analisis sentimen membantu memahami respons masyarakat terhadap berbagai isu dengan mengelompokkan opini ke dalam kategori positif, negatif, dan netral. Metode ini dapat diterapkan pada beragam domain seperti isu sosial, politik, maupun ulasan produk digital, sehingga memberikan gambaran sistematis tentang persepsi publik. Dengan membaca pola opini secara objektif, analisis sentimen juga mendukung pengambilan keputusan berbasis data.

4) *Support Vector Machine* (SVM)

Sebagai algoritma supervised learning yang mampu menangani data berdimensi tinggi, SVM dikenal stabil dan akurat dalam tugas klasifikasi berbasis fitur numerik. Metode ini bekerja dengan mencari hyperplane yang memiliki margin terbesar terhadap titik data terdekat sehingga menghasilkan pemisahan kelas yang optimal. Keunggulannya semakin menonjol pada data teks berbasis TF-IDF, di mana SVM terbukti memberikan generalisasi yang kuat dan performa yang konsisten.

5) Random Forest

Random Forest menghasilkan prediksi akhir melalui voting mayoritas dari seluruh decision tree, yang masing-masing memiliki struktur berbeda karena pemilihan fitur secara acak. Sebagai bagian dari ensemble learning, metode ini

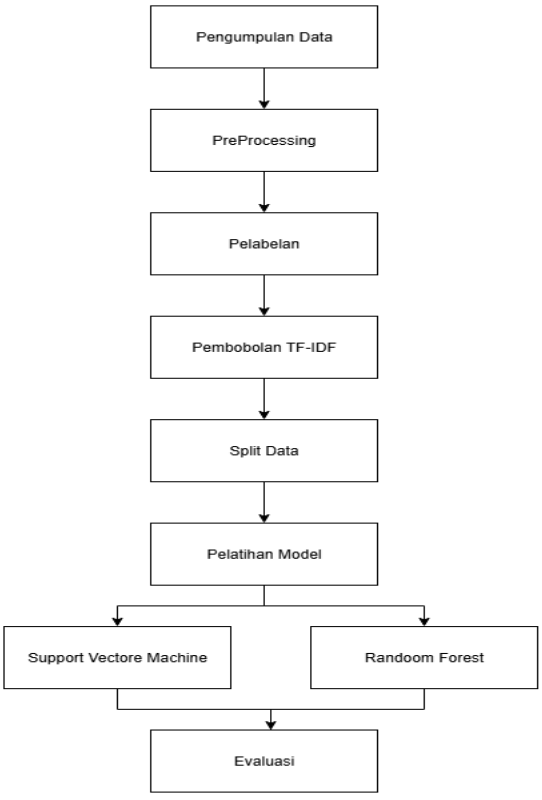
membangun banyak tree secara paralel dengan menggunakan subset fitur yang bervariasi. Keunggulannya terletak pada stabilitas yang lebih baik dan risiko overfitting yang lebih rendah dibandingkan satu decision tree. Dalam praktiknya, Random Forest telah digunakan untuk klasifikasi teks, termasuk deteksi cyberbullying di Instagram dan pengelompokan komentar ke sentimen positif, negatif, maupun netral.

6) Evaluasi Model

Evaluasi model dilakukan menggunakan empat metrik utama akurasi, presisi, recall, dan F1-score untuk memastikan performa klasifikasi tetap seimbang pada tiga kelas sentimen: positif, negatif, dan netral. Akurasi memberikan gambaran keseluruhan proporsi prediksi yang benar, sementara presisi menilai ketepatan prediksi positif, recall mengukur keberhasilan model menemukan seluruh data positif aktual, dan F1-score menyeimbangkan kedua nilai tersebut. Pendekatan ini membantu mencegah bias terhadap kelas tertentu dan memberikan penilaian kinerja yang lebih menyeluruh.

METODE PENELITIAN

Penelitian ini membandingkan kinerja Support Vector Machine (SVM) dan Random Forest (RF) dalam mengklasifikasikan sentimen pada komentar YouTube mengenai hak veto Amerika Serikat di PBB. Melalui pendekatan yang sistematis, terstruktur, dan komprehensif, proses analisis mencakup pengumpulan data, pra-pemrosesan, pelabelan sentimen, ekstraksi fitur, pembagian data, pelatihan model, hingga evaluasi performa. Setiap tahap dirancang saling terhubung secara logis sehingga hasil klasifikasi yang diperoleh akurat dan dapat dipertanggungjawabkan.



Gambar 1. Metodologi Penelitian

Pengumpulan Data

Penelitian ini menggunakan volume data sebesar 5.000 komentar YouTube yang dikumpulkan melalui teknik web scraping untuk mengotomatisasi proses pengambilan data. Fokus utama komentar tersebut adalah diskusi mengenai hak veto Amerika Serikat di PBB, sehingga YouTube dipilih sebagai sumber data karena berfungsi sebagai ruang publik tempat pengguna menyampaikan opini mereka secara langsung dan spontan.

Preprocessing

Tahapan *preprocessing* yang dilakukan dalam penelitian ini meliputi beberapa proses berikut:

a. *Data Cleaning*: Menghilangkan elemen yang tidak relevan seperti tanda baca (dalam hal ini, elipsis "...").

Contoh: "hanya nuklir yang akan menghentikan kekejaman ..." → "hanya nuklir yang akan menghentikan kekejaman"

b. Case Folding: Mengubah seluruh huruf menjadi huruf kecil untuk menjaga konsistensi.

Contoh: "Hanya nuklir yang akan menghentikan kekejaman ..." → "hanya nuklir yang akan menghentikan kekejaman ..."

c. Tokenizing: Memecah kalimat yang sudah bersih menjadi unit-unit kata individual (token).

Contoh: "hanya nuklir yang akan menghentikan kekejaman" → ['hanya', 'nuklir', 'yang', 'akan', 'menghentikan', 'kekejaman']

d. Stopwords Removal: Menghapus katakata umum yang tidak memiliki makna sentimen, seperti "hanya", "yang", dan "akan".

Contoh: ['hanya', 'nuklir', 'yang', 'akan', 'menghentikan', 'kekejaman'] → ['nuklir', 'menghentikan', 'kekejaman']

e. Stemming: Mengubah setiap kata berimbuhan ke bentuk dasarnya. Kata "menghentikan" menjadi "henti" dan "kekejaman" menjadi "kejam".

Contoh: ['nuklir', 'menghentikan', 'kekejaman'] → ['nuklir', 'henti', 'kejam']

f. Penggabungan kalimat: Menggabungkan kembali semua token yang telah diolah menjadi satu kalimat utuh.

Contoh: ['nuklir', 'henti', 'kejam'] → "nuklir henti kejam"

Pelabelan

Penentuan label sentimen dilakukan dengan menelaah makna kontekstual setiap komentar, khususnya yang berkaitan dengan topik penelitian mengenai hak veto AS di PBB. Analisis dilakukan dengan mempertimbangkan nuansa dan konteks pembahasan agar relevan dengan isu politik internasional yang dikaji. Proses klasifikasi kemudian memetakan komentar ke dalam dua kategori sentiment positif dan negative dan seluruh tahapan ini dilaksanakan setelah pra-pemrosesan data selesai dilakukan.

Pembobolan TF-IDF

TF-IDF digunakan untuk memberi bobot pada kata berdasarkan tingkat kepentingannya dalam komentar serta sebarannya di seluruh dokumen, sehingga setiap teks dapat direpresentasikan sebagai vektor numerik. Melalui transformasi ini, komentar berubah menjadi data yang siap diproses oleh algoritma klasifikasi seperti SVM dan Random Forest. Tujuan utamanya adalah mengonversi teks ke bentuk numerik agar dapat digunakan dalam tahap pemodelan machine learning.

Split Data

Pembagian data sebesar 80% untuk pelatihan dan 20% untuk pengujian digunakan sebagai praktik umum dalam machine learning untuk memastikan akurasi prediksi pada data yang belum pernah dilihat. Data latih berfungsi membangun model, sedangkan data uji digunakan untuk menilai performanya pada kondisi riil. Skema ini membantu menghasilkan evaluasi yang objektif sekaligus mengukur kemampuan generalisasi model.

Klasifikasi atau Pelatihan Model

Dua algoritma klasifikasi yang berbeda dilatih menggunakan data latih:

a. ***Support Vector Machine*** **(SVM):** SVM mencari hyperplane dengan margin maksimal untuk memisahkan sentimen, sehingga menghasilkan batas keputusan yang optimal pada ruang fitur berdimensi tinggi. Metode ini cocok untuk data teks berbasis TF-IDF karena mampu menangani banyak fitur numerik, dan terbukti memberikan akurasi tinggi, performa stabil, serta generalisasi yang baik.

b. ***Random Forest*** **(RF)**: Random Forest menggabungkan prediksi banyak decision tree melalui voting mayoritas, sehingga keputusan lebih stabil dan tidak bergantung pada satu tree. Dengan melatih setiap tree pada sampel dan fitur berbeda, algoritma ensemble ini menciptakan keragaman yang meningkatkan akurasi dan menghasilkan klasifikasi yang lebih komprehensif.

Evaluasi

Tahap akhir penelitian adalah mengevaluasi performa model SVM dan Random Forest menggunakan data uji. Evaluasi dilakukan dengan empat metrik dari confusion matrix *accuracy*, *precision*, *recall*, dan F1-*score* untuk menilai ketepatan model dalam mengklasifikasikan sentimen komentar YouTube. Dengan metrik yang sama untuk kedua model, kinerja SVM dan Random Forest dapat dibandingkan secara langsung dan objektif.

PEMBAHASAN

Hasil dan pembahasan untuk Analisis Sentimen yang sudah dilakukan yakni sebagai berikut.

Pengumpulan Data

Pengumpulan data dilakukan melalui scraping menggunakan YouTube API, dengan kata kunci “Amerika Serikat veto PBB” untuk mendapatkan video yang relevan. Proses crawling dibatasi maksimal 20 video, dan dari tiap video diambil hingga 5.000 komentar. Hasil akhirnya diperoleh 3.363 komentar yang memenuhi kriteria, lebih sedikit dari kapasitas maksimal karena tidak semua video memiliki jumlah komentar penuh atau karena adanya proses penyaringan relevansi.

```
Jumlah komentar yang sudah terkumpul: 3363
Berhasil mengumpulkan 3363 komentar.
                                                Review
0  Ngapain ngurusin veto ,yang penting besok masi...
1                                               Mantap
2  5 negara pemilik hak peto di PBB .\n1 Amerika ...
3  Kok paradok logikanya negara pemilik.hak veto ...
4  Gampang supaya as mau mengakui pelestina merde...
```

Gambar 2. Hasil *Scraping*

Preprocessing

Proses pra-pemrosesan dilakukan secara berurutan dan sistematis untuk mengubah data teks mentah hasil scraping menjadi lebih terstruktur sehingga siap

digunakan oleh algoritma. Tahapan yang dilalui mencakup pembersihan karakter tidak relevan, penyamaan huruf menjadi lowercase melalui case folding, normalisasi kata tidak baku, pemecahan teks menjadi token, penghapusan kata umum yang tidak informatif, serta stemming untuk mengembalikan kata ke bentuk dasarnya. Seluruh langkah ini bertujuan menata data teks tidak terstruktur agar dapat diolah secara efektif dalam tahapan analisis berikutnya.

	text_clean
0	Ngapain ngurusin veto yang penting besok masih...
1	Mantap
2	negara pemilik hak peto di PBB Amerika Serik...
3	Kok paradok logikanya negara pemilikhak veto p...
4	Gampang supaya as mau mengakui pelestina merde...

Gambar 3. Hasil *Cleaning*

Normalisasi

Normalisasi dilakukan setelah case folding dan sebelum tokenizing, dengan menggunakan kamus kata baku untuk memetakan kata informal atau slang menjadi bentuk standar. Tahap ini memastikan setiap kata telah dibakukan sehingga proses selanjutnya lebih konsisten.

	text_normalisasi
0	mengapai mengurusi veto yang penting besok mas...
1	mantap
2	negara pemilik hak peto di pbb amerika serikat...
3	kok paradok logikanya negara pemilikhak veto p...
4	gampang supaya as mau mengakui pelestina merde...

Gambar 5. Hasil Normalisasi

Tokenizing

Tokenizing memecah komentar menjadi kata-kata individual sehingga menghasilkan daftar token yang lebih terstruktur. Tahap ini berfungsi menyiapkan data untuk proses berikutnya, seperti stopword removal dan stemming, dengan menyediakan kumpulan kata yang siap dibersihkan lebih lanjut.

	text_tokenizing
0	[mengapai, mengurusi, veto, yang, penting, bes...
1	[mantap]
2	[negara, pemilik, hak, peto, di, pbb, amerika,...
3	[kok, paradok, logikanya, negara, pemilikhak, ...
4	[gampang, supaya, as, mau, mengakui, pelestina...

Gambar 6. Hasil *Tokenizing*

Stopword Removal

Stopword removal bertujuan menghapus kata-kata yang tidak signifikan agar "noise" dalam data teks berkurang. Dengan menyisakan kata-kata penting saja, proses ini menyederhanakan teks dan membantu model machine learning bekerja lebih efisien, sehingga makna relevan lebih mudah ditangkap dan performa klasifikasi sentimen menjadi lebih optimal.

	text_stopword
0	[mengapai, mengurusi, veto, besok, beli, beras...
1	[mantap]
2	[negara, pemilik, hak, peto, pbb, amerika, ser...
3	[paradok, logikanya, negara, pemilikhak, veto,...
4	[gampang, mengakui, pelestina, merdeka, kirim,...

Gambar 7. Hasil *Stopword Removal*

Pelabelan

Pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan berbasis kata kunci melalui kamus keywords yang telah disusun. Kata seperti *salah*, *buruk*, *parah*, *gagal*, *tidak*, *berbahaya*, dan *boikot* digunakan sebagai indikator sentimen negatif, sementara *bagus*, *setuju*, *mantap*, *benar*, *solusi*, *dukung*, dan *baik* menjadi penanda sentimen positif. Dengan metode ini, setiap komentar berhasil

diberi label secara otomatis dan siap masuk ke tahap analisis serta klasifikasi berikutnya.

	text_stemming	Sentimen
1	mantap	positif
6	amerika negaranegara dukung palestina	positif
8	palestina merdeka dukung negara negara sahabat...	positif
24	mantap dukung palestina satu mdh mdhan lindung...	positif
33	moga geeran duluan aku palestina merdeka israe...	positif

Gambar 8. Hasil Pelabelan

Pembobotan TF-IDF

Transformasi teks dilakukan untuk mengubah komentar menjadi representasi numerik agar dapat diproses oleh algoritma machine learning. Bobot setiap kata dihitung berdasarkan frekuensi kemunculannya dalam komentar serta penyebarannya di seluruh korpus. Hasilnya berupa matriks berukuran 3363 × 5000, dengan 3363 dokumen dan 5000 fitur kata unik. Melalui konversi ini, data teks menjadi vektor numerik yang siap digunakan dalam tahap pelatihan model klasifikasi.

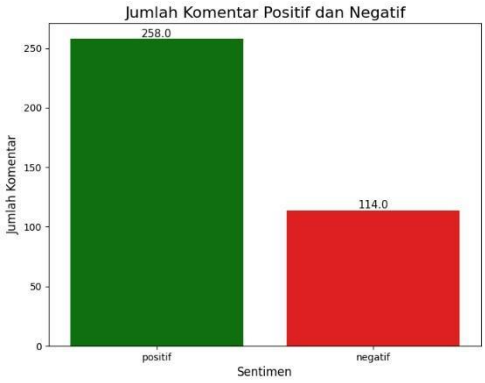
```
Proses pembobotan TF-IDF selesai.
Ukuran matriks TF-IDF: (3363, 5000)

Contoh matriks bobot TF-IDF:
   aba  abad  abadi  abai  abal  abang  abcdefghiyk  abdul  abdulwati  abiizz
0  0.0   0.0    0.0   0.0   0.0    0.0          0.0    0.0        0.0     0.0
1  0.0   0.0    0.0   0.0   0.0    0.0          0.0    0.0        0.0     0.0
2  0.0   0.0    0.0   0.0   0.0    0.0          0.0    0.0        0.0     0.0
3  0.0   0.0    0.0   0.0   0.0    0.0          0.0    0.0        0.0     0.0
4  0.0   0.0    0.0   0.0   0.0    0.0          0.0    0.0        0.0     0.0
```

Gambar 9. Hasil Pembobotan TF-IDF

Perbandingan Kelas

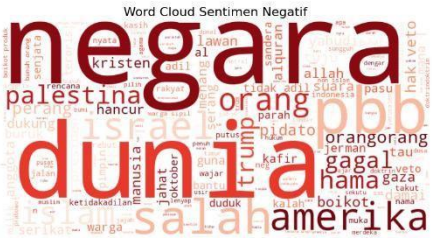
Dari hasil Pelabelan otomatis data yang didapat positif sebanyak 258 dan negatif sebanyak 114 ulasan kemudian berikut Adalah perbandingan antara kedua sentiment tersebut



Gambar 10. Hasil Perbandingan Kelas

Visualisasi Sentimen Negatif

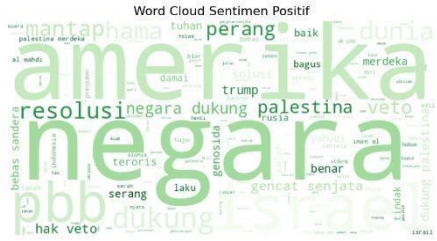
Berdasarkan hasil tahapan *preprocessing*, berikut adalah melakukan visualisasi terhadap sentimen negatif untuk memahami pola dan karakteristik komentar pada kategori tersebut.



Gambar 11. Sentiment Negatif

Visualisasi Sentimen Positif

Berdasarkan hasil tahapan *preprocessing* yang telah diselesaikan, visualisasi berikut menggambarkan distribusi sentimen positif pada data komentar.



Gambar 12. Sentimen Positif

Pelatihan Model

Berikut pelatihan model yang dilakukan menggunakan Algoritma *Support Vector Machine* dengan kernel Linear dan Algoritma *Random Forest*.

Support Vector Machine

Berikut hasil dari pelatihan model menggunakan kernel sigmoid, didapatkan akurasi 0.88 % berikut confusion matrixnya.

```
Model: SVM Sigmoid
Akurasi: 0.8800

Classification Report:
              precision    recall  f1-score   support

     negatif     0.8500    0.7391    0.7907        23
     positif     0.8909    0.9423    0.9159        52

    accuracy                         0.8800        75
   macro avg     0.8705    0.8407    0.8533        75
weighted avg     0.8784    0.8800    0.8775        75
```

Gambar 13. Hasil SVM

Random Forest

Berikut hasil dati pelatihan model dengan algoritma *Random forest,* didapatkan akurasi 0.89, berikut confusion matrixnya.

```
Model: Random Forest
Akurasi: 0.8933

Classification Report:
              precision    recall  f1-score   support

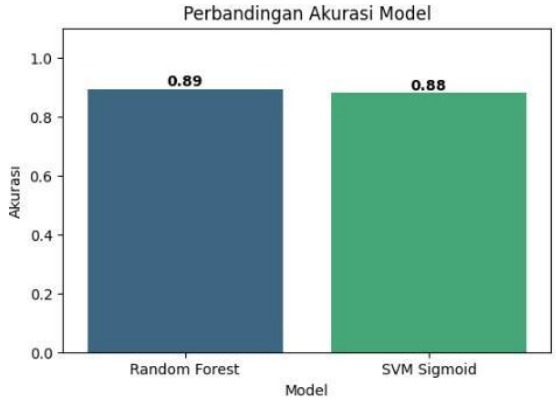
     negatif     0.9412    0.6957    0.8000        23
     positif     0.8793    0.9808    0.9273        52

    accuracy                         0.8933        75
   macro avg     0.9102    0.8382    0.8636        75
weighted avg     0.8983    0.8933    0.8882        75
```

Gambar 14. Hasil *Random Forest*

Visualisasi Perbandingan SVM dan Random Forest

Dari hasil pelatihan model berikut perbandingan akurasi model dalam pelatihan menunjukan bahwa algoritma Random forest 89% lebih unggul daripada SVM 88%.



Gambar 15. Perbandingan akurasi

KESIMPULAN

Random Forest mencapai akurasi tertinggi sebesar 89,33%, sedangkan SVM dengan kernel sigmoid memperoleh akurasi 88,00%. Random Forest lebih unggul dalam akurasi keseluruhan, sementara SVM menunjukkan performa yang lebih seimbang dalam mengklasifikasikan komentar positif dan negatif. Kesimpulannya, Random Forest lebih efektif bila tujuan utama adalah akurasi global, sedangkan SVM lebih cocok digunakan ketika keseimbangan klasifikasi antar kelas menjadi prioritas utama.

Saran

Penelitian selanjutnya dapat memperluas cakupan data dengan memasukkan isu-isu politik global lain untuk menguji kemampuan generalisasi model. Metode representasi teks yang lebih modern seperti Word2Vec atau BERT juga dapat dieksplorasi agar konteks makna kata tertangkap lebih baik dibandingkan TF-IDF. Selain itu, performa SVM dan Random Forest masih dapat ditingkatkan melalui optimasi hyperparameter yang lebih mendalam, sehingga model bekerja lebih optimal dan akurat.

DAFTAR PUSTAKA

Afdhal. (2022). *Penerapan algoritma Random Forest untuk analisis sentimen komentar di YouTube tentang Islamofobia*. Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI).

Alek. (2020). *Analisis sentimen pada media sosial Twitter menggunakan metode Naïve Bayes*. Jurnal Informatika, 7(2), 112–120.

Devi, D. L., Arifiyanti, A. A., & Wati, S. F. A. (2024). *Analisis sentimen ulasan pengguna Access by KAI menggunakan metode Word2Vec dan algoritma SVM*. Jurnal Informatika dan Teknik Elektro Terapan (JITET), 12(3).

Fadhillah, O. S. D., Jaman, J. H., & Carudin, C. (2025). *Perbandingan Naive Bayes, Support Vector Machine, Logistic Regression dan Random Forest dalam analisis sentimen mengenai TikTokShop*. Jurnal Informatika dan Teknik Elektro Terapan (JITET), 13(1).

Fastabiqul Khusna, F., Khothibul Umam, K., Siti Nur'aini, S., & Handayani, M. R. (2024). *Evaluasi efektivitas Support Vector Machine dan Random Forest dalam klasifikasi ulasan pengguna aplikasi streaming Vidio*. JSiI (Jurnal Sistem Informasi Interdisipliner).

Handayani, F., & Sari, M. (2020). *Analisis sentimen media sosial terhadap isu publik menggunakan metode machine learning*. Jurnal Sistem Cerdas, 8(3), 145–154.

Kurniawan, A., & Hidayat, R. (2021). *Evaluasi model machine learning untuk analisis sentimen dengan confusion matrix*. Jurnal Teknologi dan Sistem Komputer, 9(3), 210–217.

Lestari, D., & Nugroho, A. (2022). *Penerapan Support Vector Machine untuk analisis sentimen ulasan produk e-commerce*. Jurnal Teknologi Informasi dan Terapan, 19(2), 87–96.

Nugroho, Y., & Santoso, D. (2020). *Random Forest untuk klasifikasi sentimen pada data komentar media sosial*. Jurnal Sains Komputer dan Informatika, 8(1), 25–32.

Prasetyo, A., & Wibowo, H. (2021). *Perbandingan algoritma machine learning untuk analisis sentimen pada komentar YouTube*. Jurnal Teknologi dan Sistem Informasi, 9(1), 55–64.

Putra, R. A., & Siregar, H. (2021). *Analisis sentimen pada media sosial Twitter menggunakan metode Support Vector Machine*. Jurnal Teknologi dan Ilmu Komputer, 18(1), 33–40.

Putri, R. D., Utomo, A. P., & Santosa, B. (2023). *Analisis sentimen review aplikasi mobile menggunakan metode klasifikasi machine learning*. Jurnal Sistem Informasi, 19(2).

Ruziqiana, M. F. (2024). *Detection of cyberbullying using SVM, Naive Bayes, and Random Forest algorithm*. Jurnal Informatika dan Teknik Elektro Terapan (JITET), 12(4).

Sari, L., & Pratama, B. (2023). *Analisis sentimen menggunakan machine learning pada data komentar aplikasi digital*. Jurnal Teknologi Informasi, 11(2), 77–86.

Setiawan, B., & Rahmawati, I. (2022). *Penerapan Support Vector Machine untuk analisis sentimen berita online*. Jurnal Informatika dan Sains, 10(2), 101–109.

Utari, E. L., & Wibowo, S. H. (2023). *Analisis komparatif algoritma SVM, Naive Bayes dan LSTM pada sentimen komentar lagu Labour*. Jurnal Informatika Teknologi dan Sains (Jinteks).