PENERAPAN SUPPORT VECTOR MACHINE (SVM) UNTUK DETEKSI HOAKS PADA BERITA ONLINE

Rifky Fachuzi, Raditia Vindua
Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang Teknik Informatika
Universitas Pamulang, Jl. Puspitek, Buaran, Kec. Pamulang, Kota Tangerang Selatan,
Banten 15310 Telp/Fax : (021) 7412566
rifkyfchzi@gmail.com

ABSTRAK

Maraknya penyebaran berita hoaks di media digital telah memicu disinformasi yang berdampak negatif pada stabilitas sosial. Namun, proses verifikasi manual memiliki keterbatasan dalam hal kecepatan dan skalabilitas. Penelitian ini bertujuan untuk menerapkan algoritma *Support Vector Machine* (SVM) dalam membangun model klasifikasi berita *online* berbahasa Indonesia secara otomatis. Data yang digunakan berjumlah 3.000 berita yang bersumber dari situs pemeriksa fakta dan media terpercaya. Metodologi penelitian meliputi *preprocessing* teks dengan NLP, ekstraksi fitur menggunakan TF-IDF, serta pemodelan SVM dengan kernel Linear dan RBF. Hasil penelitian menunjukkan bahwa SVM kernel Linear mencapai performa terbaik dengan akurasi 99,8%, presisi 99,6%, dan *recall* 100%. Hal ini membuktikan bahwa metode SVM sangat efektif untuk deteksi hoaks otomatis dan dapat dikembangkan lebih lanjut menjadi sistem aplikasi.

Kata kunci : *Hoaks, SVM, klasifikasi teks, TF-IDF, NLP*

1. PENDAHULUAN

Perkembangan teknologi digital dan internet telah mengubah paradigma distribusi informasi secara fundamental, memungkinkan data tersebar luas tanpa batasan ruang dan waktu. Namun, kemudahan ini menjadi pedang bermata dua karena internet juga menjadi lahan subur bagi penyebaran hoaks atau disinformasi. Fenomena ini membawa dampak destruktif yang signifikan, mulai dari memicu kepanikan publik dan konflik sosial hingga menurunkan tingkat kepercayaan masyarakat terhadap institusi resmi serta media massa. Dalam ekosistem digital yang sangat cepat, hoaks yang tidak terkendali dapat mengancam stabilitas psikologis masyarakat melalui kecemasan dan kesalahpahaman informasi yang masif [1] [2].

Upaya mitigasi penyebaran hoaks saat ini masih sangat bergantung pada metode konvensional, yaitu verifikasi fakta secara manual oleh tenaga ahli. Meskipun memiliki akurasi yang baik, metode manual ini memiliki keterbatasan kritis dalam aspek kecepatan dan skalabilitas, sehingga tidak mampu mengimbangi volume berita digital yang diproduksi setiap detiknya.

Diperlukan sebuah pendekatan yang lebih dinamis dan otomatis untuk mengidentifikasi pola-pola disinformasi dalam skala besar guna menekan dampak negatif sebelum berita tersebut meluas. [3].

Penelitian ini bertujuan untuk mengatasi celah tersebut dengan membangun sebuah model klasifikasi otomatis berbasis *Machine Learning* untuk mendeteksi berita hoaks berbahasa Indonesia. Fokus utama penelitian adalah menerapkan algoritma *Support Vector Machine* (SVM) untuk memisahkan kategori berita antara fakta dan hoaks secara presisi. Melalui penelitian ini, diharapkan performa model dapat dievaluasi secara mendalam menggunakan metrik *accuracy*, *precision*, dan *recall* untuk memastikan keandalan sistem dalam aplikasi dunia nyata.

Untuk mencapai tujuan tersebut, penelitian ini mengusulkan alur penyelesaian masalah berbasis *Natural Language Processing* (NLP) dan algoritma SVM. Langkah awal melibatkan tahapan preprocessing teks yang ketat (seperti *case folding, cleaning, tokenization, stopword removal, dan stemming*) untuk menghilangkan noise yang dapat mengganggu performa model.

Selanjutnya, data teks dikonversi menjadi representasi numerik menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) untuk mengekstraksi bobot kata yang paling relevan sebagai fitur pembeda [4].

Model kemudian diklasifikasikan menggunakan algoritma SVM yang bekerja dengan mencari hyperplane optimal untuk memisahkan data ke dalam dua kelas dengan margin maksimal [5].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun model klasifikasi hoaks menggunakan algoritma SVM serta mengevaluasi performanya menggunakan metrik *accuracy*, *precision*, dan *recall*.

2. TINJAUAN PUSTAKA

2.1. Hoaks

Hoaks merupakan informasi yang direkayasa secara sengaja untuk menutupi informasi sebenarnya atau memanipulasi opini publik dengan tujuan menyesatkan [6].

Di era digital, penyebaran hoaks meningkat secara eksponensial karena karakteristik internet yang memungkinkan distribusi konten tanpa melalui proses verifikasi yang ketat seperti pada media konvensional. Hoaks sering kali dirancang dengan judul yang provokatif dan emosional untuk memicu reaksi cepat

dari pembaca, sehingga penyebarannya di media sosial terjadi secara masif dalam waktu singkat. [1].

2.2. *Machine Learning*

Machine learning adalah disiplin ilmu dalam kecerdasan buatan yang berfokus pada pengembangan algoritma yang memungkinkan komputer untuk belajar dari data secara mandiri tanpa pemrograman eksplisit. Dalam klasifikasi teks, *machine learning* bekerja dengan mengenali pola statistik dari kumpulan data latih untuk membangun model prediktif. Terdapat dua kategori utama dalam metode ini, yaitu *supervised learning* yang menggunakan data berlabel (seperti pada penelitian ini) dan *unsupervised learning* yang bekerja pada data tanpa label.[7].

2.3. *Natural Language Processing* (NLP)

Natural Language Processing (NLP) adalah bidang teknologi yang menjembatani interaksi antara bahasa manusia dengan sistem komputer. Karena komputer hanya dapat memproses data numerik, NLP berperan penting dalam mentransformasi teks mentah yang tidak terstruktur menjadi format yang dapat diolah [8].

Tahapan *preprocessing* dalam NLP sangat krusial karena teks berita sering mengandung elemen yang tidak relevan yang dapat menurunkan akurasi klasifikasi. Tahapan tersebut meliputi :

a. *Case folding*, yaitu Penyeragaman seluruh karakter menjadi huruf kecil agar sistem tidak menganggap kata "Hoaks" dan "hoaks" sebagai dua entitas yang berbeda
b. Pembersihan karakter, yaitu Penghapusan simbol, angka, dan tanda baca yang tidak memberikan informasi kontekstual dalam pendeteksian berita Tokenisasi, untuk memecah kalimat menjadi kata-kata
c. *Stopword removal*, yaitu Proses eliminasi kata-kata umum (seperti "dan", "yang", "di") yang muncul sangat sering namun tidak memiliki makna signifikan dalam membedakan kategori berita.
d. *Stemming*, yaitu Pengubahan kata berimbuhan menjadi kata dasarnya (misalnya "menyebarkan" menjadi "sebar") untuk mengurangi dimensi fitur dan menjaga konsistensi makna.

2.4. *Term Frequency–Inverse Document Frequency* (TF-IDF)

TF-IDF merupakan metode ekstraksi fitur yang digunakan untuk mengubah data teks menjadi bentuk numerik sehingga dapat diproses oleh algoritma *machine learning* [9].

Metode ini bekerja dengan menghitung bobot setiap kata berdasarkan dua komponen utama, yaitu *Term Frequency* (TF) yaitu Menghitung seberapa sering suatu kata muncul dalam sebuah dokumen tertentu. Dan *Inverse Document Frequency* (IDF) yaitu Mengukur seberapa langka atau unik suatu kata di seluruh kumpulan dokumen (dataset). Dengan mengalikan nilai TF dan IDF, kata-kata yang menjadi ciri khas unik sebuah berita hoaks akan mendapatkan bobot yang lebih tinggi, sementara kata umum yang muncul di semua berita akan mendapatkan bobot yang rendah.

Dengan menggunakan TF-IDF, kata-kata yang sering muncul dalam satu dokumen tetapi jarang muncul di dokumen lain akan memiliki bobot tinggi. Hal ini membantu model dalam membedakan karakteristik antara berita hoaks dan fakta.

2.5. *Support Vector Machine* (SVM)

Support Vector Machine (SVM) adalah algoritma supervised learning yang sangat efektif untuk klasifikasi teks berdimensi tinggi. Prinsip kerja utama SVM adalah mencari *hyperplane* (bidang pemisah) optimal yang memisahkan dua kelas dengan margin maksimal [4].

SVM bekerja dengan menentukan batas pemisah (*hyperplane*) yang memiliki margin terbesar antara dua kelas. Titik data yang berada paling dekat dengan *hyperplane* disebut *support vector*, dan memiliki peran penting dalam menentukan posisi *hyperplane* tersebut.

Keunggulan SVM adalah kemampuannya dalam menangani data berdimensi tinggi seperti data teks hasil TF-IDF, serta kemampuannya dalam menghasilkan model dengan generalisasi yang baik. SVM juga banyak digunakan dalam berbagai penelitian klasifikasi teks karena kemampuannya dalam menangani data berdimensi tinggi [10].

3. METODE PENELITIAN

3.1. *Dataset*

Dataset yang digunakan dalam penelitian ini terdiri dari 3.000 berita *online* berbahasa Indonesia yang terbagi menjadi dua kategori, yaitu berita hoaks dan berita fakta. *Dataset* tersebut terdiri dari 2.000 data latih dan 1.000 data uji.

Data berita hoaks diperoleh dari situs Turnbackhoax, sedangkan data berita fakta diperoleh dari portal berita terpercaya seperti Kompas dan Detik. Pembagian *dataset* ini dilakukan untuk memastikan model dapat dilatih dan diuji secara optimal.

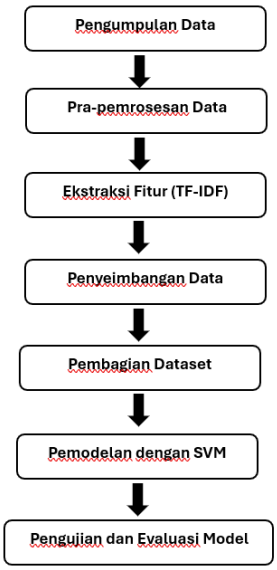
Tabel 1. *Dataset*

Nama File	Jumlah Data	Jenis Data	Keterangan
DataBerita.csv	614	Campuran	Berita umum (hoaks dan fakta)
turnbackhoax.csv	3.071	Hoaks	*Dataset* hasil verifikasi TurnBackHoax.id
berita.csv	541	Fakta	Berita aktual dari portal berita terpercaya
DataDetik.csv	4.945	Fakta	Artikel lain dari situs Detik.com

Berdasarkan tabel tersebut, dapat dilihat bahwa distribusi data sebelum penyeimbangan belum merata antara kelas hoaks dan fakta. Ketidakseimbangan ini

berpotensi menyebabkan model menjadi bias terhadap salah satu kelas. Oleh karena itu, dilakukan proses penyeimbangan data untuk meningkatkan performa model dalam proses klasifikasi.

3.2. Alur Penelitian



Gambar 1. Kerangka Kerja Penelitian

Penelitian ini dilakukan melalui beberapa tahapan utama yang disusun secara sistematis mulai dari pengumpulan data hingga evaluasi model. Proses dimulai dengan pengumpulan data, kemudian dilakukan *preprocessing* untuk membersihkan data teks. Selanjutnya dilakukan ekstraksi fitur menggunakan TF-IDF, dilanjutkan dengan penyeimbangan data, pembagian *dataset*, pemodelan menggunakan SVM, dan terakhir evaluasi model.

Berdasarkan Gambar 1, dapat dilihat bahwa setiap tahapan saling berkaitan dan membentuk alur kerja yang terstruktur dalam proses klasifikasi berita hoaks.

3.3. *Preprocessing*

Preprocessing merupakan tahap penting dalam pengolahan data teks untuk meningkatkan kualitas data sebelum dilakukan pemodelan. Tahapan *preprocessing* yang dilakukan adalah sebagai berikut:

a. *Case Folding*
Proses ini bertujuan untuk mengubah seluruh teks menjadi huruf kecil agar tidak terjadi perbedaan makna akibat perbedaan kapitalisasi.

Tabel 2. Contoh *Case Folding*

Teks asli	"Berita Ini HOAKS dan Jangan Sebar Lagi!"
Hasil	"berita ini hoaks dan jangan sebar lagi"

Dari Tabel 1 dapat dilihat bahwa kata dengan huruf besar diubah menjadi huruf kecil sehingga menghasilkan bentuk teks yang seragam.

b. Pembersihan Karakter
Tahapan ini bertujuan untuk menghapus karakter yang tidak relevan seperti angka, simbol, dan tanda baca.

Tabel 3. Pembersihan Karakter

Teks asli	"COVID-19!!! Ini bukan fakta, tapi HOAKS 😡 #waspada"
Hasil	"covid ini bukan fakta tapi hoaks waspada"

Berdasarkan Tabel 3, terlihat bahwa karakter yang tidak diperlukan berhasil dihilangkan sehingga teks menjadi lebih bersih.

c. Tokenisasi
Tokenisasi adalah proses memecah kalimat menjadi kata-kata atau token.

Tabel 4. Tokenisasi

Teks asli	"berita ini hoaks dan jangan sebar lagi"
Hasil	['berita', 'ini', 'hoaks', 'dan', 'jangan', 'sebar', 'lagi']

Dari Tabel 4 dapat dilihat bahwa kalimat dipecah menjadi beberapa kata yang siap untuk diproses lebih lanjut.

d. *Stopword Removal*
Stopword removal bertujuan untuk menghapus kata-kata umum yang tidak memiliki makna signifikan.

Tabel 5. *Stopword Removal*

Sebelum	['berita', 'ini', 'hoaks', 'dan', 'jangan', 'sebar', 'lagi']
Sesudah	['berita', 'hoaks', 'jangan', 'sebar']

Berdasarkan Tabel 5 kata-kata yang tidak penting berhasil dihapus sehingga hanya tersisa kata-kata utama.

e. Penggabungan Teks
Setelah proses pembersihan selesai, token yang tersisa digabung kembali menjadi kalimat.

Tabel 6. Penggabungan Teks

Sebelum	[‘berita’, ‘hoaks’, ‘jangan’, ‘sebar’]
Sesudah	“berita hoaks jangan sebar”

Dari Tabel 6 dapat dilihat bahwa teks yang dihasilkan menjadi lebih ringkas dan siap untuk proses ekstraksi fitur.

3.4. Ekstrasi Fitur

Pada tahap ini, data teks yang telah diproses diubah menjadi representasi numerik menggunakan metode TF-IDF. Proses ini menghasilkan vektor fitur yang merepresentasikan setiap dokumen dalam bentuk angka sehingga dapat diproses oleh algoritma SVM.

Proses ekstraksi fitur menggunakan TF-IDF memiliki peran penting dalam meningkatkan kualitas representasi data teks. Dengan metode ini, setiap kata diberikan bobot berdasarkan tingkat kepentingannya dalam dokumen, sehingga kata-kata yang relevan

memiliki kontribusi lebih besar dalam proses klasifikasi.

Selain itu, penggunaan TF-IDF mampu mengurangi pengaruh kata-kata umum yang sering muncul namun tidak memiliki makna khusus. Hal ini membantu model dalam fokus pada kata-kata yang menjadi pembeda antara berita hoaks dan fakta, sehingga meningkatkan akurasi model.

3.5. Pemodelan SVM

Pemodelan dilakukan menggunakan algoritma Support *Vector Machine* dengan beberapa konfigurasi parameter dan *kernel*. Model dilatih menggunakan data latih dan kemudian diuji menggunakan data uji untuk melihat performanya.

Tabel 7. Konfigurasi Model SVM

Parameter	*Kernel Linear*	*Kernel* RBF
Jumlah fitur TF-IDF	1000	1000
Nilai C (*Regularization*)	1.0	1.0
Pembagian data	2.000 : 1.000	2.000 : 1.000
Sratifikasi kelas	Aktif (*stratify*=y)	Aktif (*stratify*=y)
Random state	42	42

Dari tabel tersebut dapat dilihat parameter yang digunakan dalam proses pelatihan model. Parameter ini berpengaruh terhadap performa model yang dihasilkan.

Pemilihan parameter dalam model SVM sangat berpengaruh terhadap performa klasifikasi. Parameter seperti nilai C digunakan untuk mengontrol tingkat regularisasi model, yaitu keseimbangan antara margin yang besar dan kesalahan klasifikasi.

Selain itu, penggunaan *kernel* yang berbeda seperti *linear* dan RBF memungkinkan model untuk menangani pola data yang berbeda. *Kernel linear* cocok untuk data yang dapat dipisahkan secara *linier*, sedangkan kernel RBF lebih fleksibel dalam menangani data yang tidak terpisah secara *linier*.

3.6. Evaluasi Model

Evaluasi model dilakukan menggunakan *confusion matrix* untuk menghitung nilai *accuracy*, *precision*, dan *recall*. *Confusion matrix* digunakan untuk memberikan gambaran detail mengenai performa model dalam mengklasifikasikan data. Dengan menggunakan *confusion matrix*, dapat diketahui jumlah data yang diklasifikasikan dengan benar maupun salah, sehingga evaluasi model menjadi lebih komprehensif [11].

4. HASIL DAN PEMBAHASAN

4.1. *Dataset*

Tabel 8. Sumber *Dataset*

Nama File	Jumlah Data	Jenis Data	Keterangan
DataBerita.csv	614	Campuran	Berita umum (hoaks dan fakta)
turnbackhoax.csv	3.071	Hoaks	*Dataset* hasil verifikasi TurnBackHoax.id
berita.csv	541	Fakta	Berita aktual dari portal berita terpercaya
DataDetik.csv	4.945	Fakta	Artikel lain dari situs Detik.com

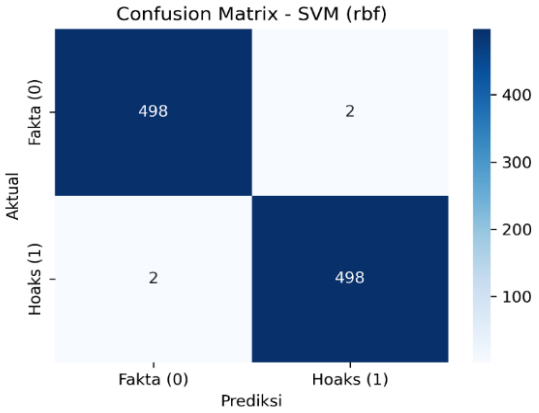
Berdasarkan tabel tersebut, *dataset* yang digunakan berasal dari beberapa sumber terpercaya sehingga dapat meningkatkan validitas penelitian.

4.2. Hasil *Preprocessing*

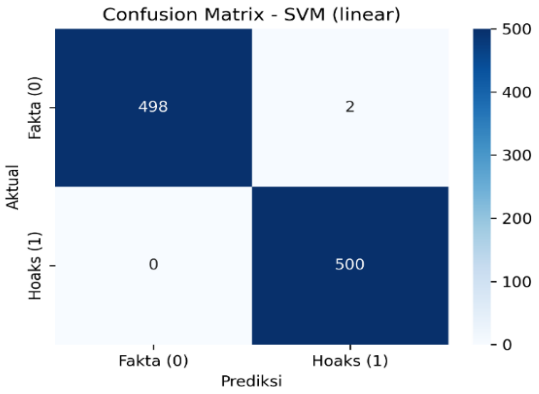
Gambar 2. Hasil *Preprocessing*

Dari gambar tersebut dapat dilihat bahwa data teks telah berhasil dibersihkan dari *noise* dan siap digunakan untuk proses klasifikasi.

4.3. Hasil Model



Gambar 3. *Confusion Matrix Model* SVM *Kernel* RBF



Gambar 4. *Confusion Matrix Model* SVM *Kernel Linear*

Berdasarkan gambar tersebut, dapat dilihat distribusi prediksi model terhadap data aktual. *Confusion matrix* menunjukkan jumlah data yang diklasifikasikan dengan benar maupun salah.

Berdasarkan hasil *confusion matrix*, dapat dilihat bahwa sebagian besar data berhasil diklasifikasikan dengan benar oleh model. Nilai *True Positive* dan *True Negative* yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali kedua kelas.

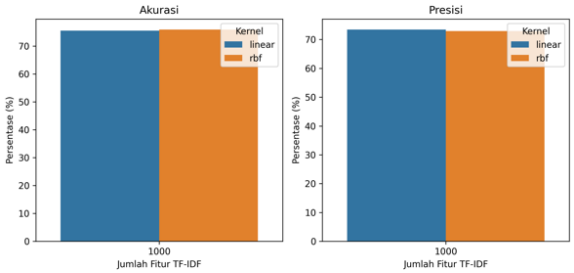
Kesalahan klasifikasi yang terjadi relatif kecil dan umumnya disebabkan oleh kemiripan struktur teks antara berita hoaks dan fakta. Hal ini menunjukkan bahwa meskipun model memiliki performa yang tinggi, masih terdapat tantangan dalam membedakan konten yang memiliki karakteristik serupa.

4.4. Evaluasi Model

Tabel 9. Hasil Evaluasi Model

Kernel	**Nilai C**	**Jumlah Fitur**	**Akurasi (%)**	**Presisi (%)**	**Recall (%)**	**F1-Score (%)**
Linear	1.0	1000	99.8	99.6	100.0	99.8
RBF	1.0	1000	99.6	99.6	99.6	99.6

Dari tabel tersebut dapat dilihat nilai *accuracy*, *precision*, dan *recall* yang menunjukkan bahwa model memiliki performa yang baik dalam mendeteksi berita hoaks.



Gambar 5. Grafik Perbandingan Akurasi dan Presisi Model SVM

Berdasarkan grafik pada Gambar 5, dapat dilihat perbandingan performa model SVM berdasarkan metrik evaluasi. Grafik menunjukkan bahwa model dengan kernel linear memiliki nilai akurasi yang sedikit lebih tinggi dibandingkan kernel RBF. Hal ini menunjukkan bahwa kernel linear lebih optimal dalam menangani data pada penelitian ini.

4.5. Analisis

Berdasarkan hasil evaluasi, model SVM menunjukkan kemampuan yang baik dalam membedakan antara berita hoaks dan fakta. Hal ini disebabkan oleh penggunaan metode TF-IDF yang mampu merepresentasikan kata-kata penting dalam dokumen secara efektif. Selain itu, proses *preprocessing* yang dilakukan secara menyeluruh juga berkontribusi dalam meningkatkan kualitas data sehingga mempermudah proses klasifikasi.

Namun demikian, terdapat beberapa kesalahan klasifikasi yang ditunjukkan pada *confusion matrix*, yang kemungkinan disebabkan oleh kemiripan struktur kalimat antara berita hoaks dan fakta. Hal ini menunjukkan bahwa meskipun model memiliki performa yang baik, masih terdapat ruang untuk peningkatan, misalnya dengan penggunaan algoritma lain atau teknik *deep learning*.

Selain itu, performa model yang tinggi menunjukkan bahwa kombinasi metode *preprocessing* dan ekstraksi fitur TF-IDF mampu menghasilkan representasi data yang baik. Hal ini memungkinkan algoritma SVM untuk memisahkan data hoaks dan fakta dengan lebih akurat.

Perbedaan performa antara *kernel linear* dan RBF juga menunjukkan bahwa pemilihan *kernel* berpengaruh terhadap hasil klasifikasi. *Kernel linear* cenderung lebih efektif untuk data dengan distribusi yang cukup terpisah secara *linier*, sedangkan *kernel* RBF lebih fleksibel namun membutuhkan penyesuaian parameter yang lebih optimal.

Dengan demikian, hasil penelitian ini membuktikan bahwa SVM merupakan metode yang efektif dalam klasifikasi teks, khususnya dalam deteksi berita hoaks berbahasa Indonesia.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan algoritma SVM untuk mendeteksi berita hoaks dengan performa yang baik. Model mampu mengklasifikasikan berita dengan tingkat akurasi yang tinggi.Penelitian selanjutnya disarankan untuk Menggunakan *dataset* lebih besar, Membandingkan dengan algoritma lain, Mengembangkan sistem berbasis aplikasi

DAFTAR PUSTAKA

[1] T. N. Faturohmah and T. A. Salim, "Perilaku Masyarakat Terhadap Penyebaran Hoax Selama Pandemi Covid-19 Melalui Media di Indonesia: Tinjauan Literatur Sistematis," *Tik Ilmeu J. Ilmu Perpust. dan Inf.*, vol. 6, no. 1, p. 121, 2022, doi: 10.29240/tik.v6i1.3432.

[2] F. Febriansyah and N. N. Muksin, "Fenomena Media Sosial: Antara Hoaks, Destruksi Demokrasi, Dan Ancaman Disintegrasi Bangsa," *Sebatik*, vol. 24, no. 2, pp. 193–200, 2020, doi: 10.46984/sebatik.v24i2.1091.

[3] T. Nurhaipah and Z. Ramallah, "Literasi Media Dalam Menangkal Informasi Hoaks Jelang Kontestasi Politik 2024," *Indones. J. Digit. Public Relations*, vol. 2, no. 2, p. 100, 2024, doi: 10.25124/ijdpr.v2i2.6834.

[4] I. A. Ropikoh, R. Abdulhakim, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) untuk Klasifikasi Berita Hoax Covid-19," *J. Appl. Informatics Comput.*, vol. 5, no. 1, pp. 64–73, 2021, doi: 10.30871/jaic.v5i1.3167.

[5] R. DickiPrabowo, I. Widaningrum, and J. Karaman, "Sistem Deteksi Berita Hoax Pemilu

2024 Indonesia Menggunakan Algoritma K-Nearest Neighbor (Knn) Dan Support Vector Machine (Svm)," *JIKO (Jurnal Inform. dan Komputer)*, vol. 9, no. 1, p. 93, 2025, doi: 10.26798/jiko.v9i1.1424.

[6] C. Juditha, "People Behavior Related To The Spread Of Covid-19's Hoax," *J. Pekommas*, vol. 5, no. 2, pp. 105–116, 2020, doi: 10.30818/jpkm.2020.2050201.

[7] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.

[8] V. Kusumawardani and B. Cahyanto, "Fenomena Buzzer dan Hoax Pada Sosial Media dalam Menentukan Pilihan Politik Bagi Gen-Z pada Pilpres 2024 dalam Perspekstif Agenda Setting," *Universitas (Stuttg).*, no. 2, pp. 241–261, 2023.

[9] T. I. Alfawas and A. Rahim, "Penerapan Fitur Ekstraksi TF-IDF untuk Analisis Sentimen Ulasan Game Bus Simulator Indonesia dengan Algoritma Naive Bayes," vol. 4, pp. 3177–3193, 2024.

[10] A. Sentimen, P. Maskapai, H. C. Husada, and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM) Sentiment Analysis of Airline on Twitter Platform Using Support Vector Machine (SVM) Algorithm," vol. 10, no. 1, pp. 18–26, 2021, doi: 10.34148/teknika.v10i1.311.

[11] Indra, Agus Umar Hamdani, Suci Setiawati, Zena Dwi Mentari, and Mauridhy Hery Purnomo, "Comparison of K-NN, SVM, and Random Forest Algorithm for Detecting Hoax on Indonesian Election 2024," *J. Nas. Pendidik. Tek. Inform.*, vol. 13, no. 1, pp. 166–179, 2024, doi: 10.23887/janapati.v13i1.76079.