ANALISIS SENTIMEN TERHADAP GERAKAN BOIKOT DI X #BOYCOTT MENGGUNAKAN SUPPORT VECTOR MACHINE

Nada Firda Khofifah[1], Intan Putri Mansyur Pratama[2], Azzahra Asti Khairunnisa[3], Anggraini Puspita Sari[4*]

[1,2 3,4]Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur
[1]nadafirdakhofifah@gmail.com, [2]intanmansyur16@gmail.com, [3]azzahraasti.13589@gmail.com,
[4*]anggraini.puspita.if@upnjatim.ac.id

.

Abstrak

Media sosial, khususnya *platform* X (sebelumnya dikenal sebagai *Twitter*), telah menjadi ruang utama bagi masyarakat untuk menyuarakan pandangan terkait isu-isu sosial, politik, dan ekonomi. Salah satu fenomena menonjol adalah gerakan *boycott* yang disuarakan melalui tagar seperti #*Boycott*, sebagai bentuk protes kolektif terhadap entitas yang dianggap bertentangan dengan nilai atau kepentingan publik. Dalam konteks konflik geopolitik, seperti ketegangan yang melibatkan Israel, tagar ini sering kali menjadi pusat diskusi daring yang intens. Penelitian ini bertujuan untuk mengkaji dinamika diskusi, kecenderungan sentimen, dan emosi yang muncul dalam wacana digital terkait *boycott* terhadap produk yang terafiliasi dengan Israel. Metode yang digunakan dalam penelitian ini adalah klasifikasi sentimen menggunakan algoritma *Support Vector Machine* (SVM). Dataset yang digunakan dikumpulkan melalui proses *crawling* terhadap tagar-tagar terkait gerakan *boycott* selama bulan Januari-April 2025, yang mencakup berbagai percakapan publik di *platform* X. Hasil pengujian menunjukkan bahwa model SVM mampu mengklasifikasikan data dengan tingkat akurasi mencapai 94%, dimana data uji berhasil diklasifikasikan dengan benar ke dalam tiga kelas sentimen: pro, netral, dan kontra. Hal ini menandakan bahwa model cukup andal dalam mengidentifikasi sentimen masyarakat. Penelitian ini menunjukkan bahwa algoritma SVM efektif dalam mengidentifikasi dan mengelompokkan sentimen masyarakat terhadap gerakan *boycott* di media sosial. Penelitian ini diharapkan dapat menjadi kontribusi awal dalam pemanfaatan *machine learning* untuk menganalisis respon publik terhadap isu-isu sosial dan kemanusiaan melalui *platform* digital.

Kata kunci: media sosial, sentimen, *boycott*, *support vector machine*, TF-IDF, SMOTE

1. Pendahuluan

Media sosial telah menjadi sarana yang penting bagi masyarakat dalam berkomunikasi dan bertukar informasi pada era digital saat ini(Ray et al., 2024). Media sosial, khususnya *platform* X (sebelumnya dikenal sebagai *Twitter*), menjadi wadah utama bagi masyarakat dalam mengekspresikan pandangan terhadap isu-isu sosial, politik, dan ekonomi. Salah satu fenomena yang menonjol adalah gerakan *boycott* yang disuarakan melalui tagar seperti #*Boycott*, sebagai bentuk protes kolektif terhadap entitas yang dianggap tidak sejalan dengan nilai atau kepentingan publik.

Kasus yang menjadi perhatian global adalah konflik antara Hamas dan Israel, yang memuncak pada Oktober 2023. Konflik ini tidak hanya memicu gelombang solidaritas dari berbagai belahan dunia, termasuk Indonesia dan Malaysia, tetapi juga menyoroti berbagai dimensi seperti politik, agama, dan hak asasi manusia, di mana rakyat Palestina kerap berada dalam kondisi yang sangat tertekan dan tidak menguntungkan(Idris Siregar et al., 2024). Masyarakat di berbagai negara seperti Korea Selatan, Brasil, dan Inggris turut melakukan aksi solidaritas melalui demonstrasi, doa bersama, hingga kampanye daring(Fitri & Ulmuftia, 2025).

Respons publik di kawasan Asia Tenggara, khususnya Indonesia dan Malaysia, tampak nyata melalui seruan gerakan *boycott* terhadap produk-produk yang diasosiasikan dengan Israel atau pihak yang mendukung kebijakan zionisme, yang dikenal sebagai gerakan *Boycott*, *Divestment*, *and Sanction* (BDS). Selama puncak konflik antara Palestina dan Israel, masyarakat Indonesia memberikan respons yang kuat, isu ini banyak diungkapkan di berbagai platform media sosial, dengan X menjadi salah satu yang paling dominan dalam penyampaian sentimen dan pandangan(Susilawati et al., 2024). Selain itu, konteks lokal juga memperkuat urgensi penelitian ini. Fatwa Majelis Ulama Indonesia (MUI) Nomor 83 Tahun 2023 tentang Hukum Dukungan terhadap Perjuangan Palestina menetapkan bahwa mendukung perjuangan kemerdekaan Palestina adalah suatu kewajiban. Dalam fatwa tersebut juga disebutkan bahwa umat Islam dianjurkan untuk menghindari penggunaan produk yang mendukung penjajahan dan zionisme(Munandar et al., 2023).

Pemahaman mengenai perkembangan opini publik Indonesia terkait gerakan *boycott* tersebut diperoleh melalui pendekatan analisis sentimen terhadap percakapan warganet di *platform* X. Analisis ini bertujuan untuk mengkaji dinamika diskusi, kecenderungan sentimen, serta emosi yang muncul dalam wacana digital mengenai *boycott* produk terafiliasi dengan Israel. Metode yang digunakan dalam penelitian ini adalah *Support Vector Machine* (SVM), yaitu algoritma machine learning yang dikenal efektif untuk tugas klasifikasi, termasuk klasifikasi sentimen(Permata Aulia et al., 2021). Algoritma SVM ini bekerja dengan memanfaatkan pendekatan berbasis statistik untuk mengenali dan membedakan pola data secara optimal(Tantika & Kudus, 2022).

Hasil dari penelitian(Alifa & Alita, 2024) menyatakan bahwa isu *boycott* produk telah mulai dibahas, khususnya dalam konteks opini publik terhadap konflik global yang melibatkan Palestina dan Israel. Berdasarkan hal tersebut, ditemukan bahwa model SVM mampu mencapai tingkat akurasi sebesar 93%. Namun, setelah dianalisis lebih lanjut, performa tinggi ini cenderung disebabkan oleh dominasi prediksi pada opini positif. Hal ini mengindikasikan bahwa model belum mampu mengenali opini negatif secara optimal, yang kemungkinan besar disebabkan oleh ketidakseimbangan data atau kecenderungan model untuk lebih mengenali pola dari kelas mayoritas.

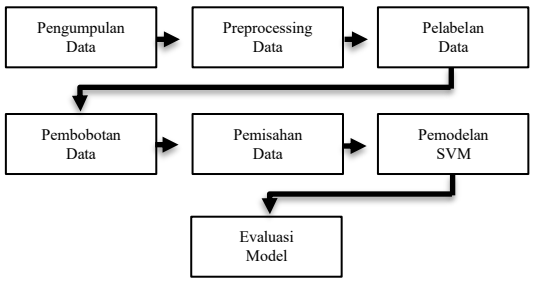
Berdasarkan Hasil dari penelitian(Tarigan & Yusupa, 2024) terkait perbandingan algoritma machine learning dalam analisis sentimen mobil listrik di Indonesia pada sosial media X didapatkan hasil bahwa algortitma SVM mendapatkan 75,62% tingkat akurasi hasil, kemudian algoritma Naive Bayes mendapatkan 60,3% tingkat akurasi hasil dan algoritma Random Forest mendapatkan 69,1% tingkat akurasi hasil. Sehingga dapat disimpulkan bahwa algoritma SVM bekerja lebih baik dibandingkan dengan algoritma naive bayes dan random forest. Temuan ini menunjukkan bahwa SVM memiliki performa klasifikasi yang lebih unggul dalam konteks pengolahan data teks, khususnya untuk analisis sentimen. Berdasarkan temuan tersebut, penelitian ini mengaplikasikan algoritma Support Vector Machine (SVM) untuk menganalisis sentimen masyarakat terhadap gerakan boikot (#Boycott) di platform X. Penelitian ini bertujuan untuk mencapai akurasi yang lebih tinggi dari penelitian sebelumnya.

Sejauh ini, jumlah penelitian yang secara spesifik menganalisis gerakan *boycott* produk menggunakan pendekatan *machine learning* masih terbatas, yaitu menggunakan SVM, studi ini berupaya mengisi celah tersebut melalui penelitian dengan pendekatan yang lebih komprehensif. Penelitian ini tidak hanya membedakan tiga kelas sentimen pro, netral, dan kontra, tetapi juga mengaitkannya langsung dengan konteks sosial-politik dari gerakan *boycott* yang berkembang di *platform* X.

Penelitian ini memberikan sejumlah kontribusi baru yang signifikan, baik dalam aspek metodologis maupun dalam hal kedalaman analisis. Pertama, penelitian ini menggunakan pelabelan sentimen dalam tiga kategori, yaitu pro (mendukung boikot), kontra (menentang boikot), dan netral, untuk menangkap kompleksitas opini publik secara lebih akurat. Kedua, berbeda dari penelitian sebelumnya yang tidak menyajikan evaluasi menyeluruh terhadap kinerja model, penelitian ini menggunakan *confusion matrix* sebagai bagian dari evaluasi performa klasifikasi, yang memberikan gambaran lebih detail terhadap kekuatan dan kelemahan model dalam mengklasifikasikan masing-masing kategori sentimen. Dengan pendekatan ini, penelitian kami diharapkan dapat memperluas cakupan literatur terkait analisis sentimen terhadap isu *boycott* produk, khususnya dalam konteks wacana publik digital di Indonesia. Selain itu, evaluasi yang lebih komprehensif ini juga dapat menjadi acuan bagi penelitian selanjutnya dalam mengembangkan model yang lebih akurat terhadap dinamika sentimen di ranah digital.

2. Metode

Penelitian ini mengikuti sejumlah tahapan yang dilakukan secara sistematis dan berurutan, sebagaimana ditampilkan pada Gambar 1(PS et al., 2022). Tahapan penelitian ini dilakukan dengan pendekatan analisis sentimen untuk mengkaji tweet pengguna X terhadap Gerakan Boikot *#Boycott*. Rancangan penelitian ini mencakup pengumpulan data, preprocessing data, pelabelan data, ekstraksi fitur, pemisahan data, dan pemodelan SVM.



Gambar 1. Diagram proses metodologi

2.1 Pengumpulan Data

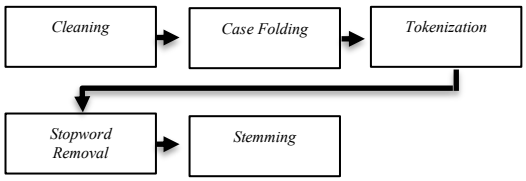
Tahap awal dalam penelitian ini adalah pengumpulan data dengan melakukan *data selection*, yaitu memilih dan mengumpulkan data yang relevan dari media sosial *Twitter* melalui proses *crawling*(Permata Aulia et al., 2021). Pengambilan data dilakukan pada bulan Januari-April 2025, dengan menggunakan kata kunci "*boycott*" untuk memperoleh data berupa komentar masyarakat terkait aksi boikot terhadap produk yang terafiliasi dengan Israel. Proses crawling dilakukan menggunakan

Twitter API yang menggunakan *twitter auth* token(Gifari et al., 2022) dan dijalankan melalui *Google Colab* dengan menggunakan bahasa pemrograman python, karena platform ini menyediakan berbagai fitur yang sangat membantu, baik bagi pemula maupun profesional, terutama dalam bidang *machine learning* dan pengolahan data(Hariyanto et al., 2020).

Hasil *crawling* disimpan dalam format file .csv dengan jumlah data 4190 data. Pada tahap ini, data yang diperoleh kemudian dianalisis untuk memahami konteks dan struktur dari *tweet*, serta mengidentifikasi jenis informasi yang terkandung di dalamnya, sebagai persiapan untuk tahap pemrosesan data lebih lanjut.

2.2 *Preprocessing* Data

Tahapan berikutnya dalam penelitian ini adalah *preprocessing* data. *Preprocessing* merupakan langkah awal sebelum melakukan klasifikasi, yang bertujuan untuk menyaring data yang tidak relevan serta menyesuaikan format teks pada data *tweet*(Alifa & Alita, 2024). *Preprocessing* juga bertujuan untuk mengolah data mentah yang telah dikumpulkan menjadi data yang bersih dan siap digunakan pada tahap selanjutnya(Aziz, 2022). Proses ini terdiri atas beberapa langkah, antara lain *cleaning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming*, yang dapat dilihat pada Gambar 2.



Gambar 2. Diagram *preprocessing* data

1. Cleaning

Tahap *cleaning* data merupakan tahap awal pada *preprocessing* data yang bertujuan untuk menghilangkan sejumlah data yang tidak lengkap(Hendrastuty et al., 2021), juga data yang tidak relevan atau tidak konsisten, sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Hasil *cleaning*

Sebelum *cleaning*	Setelah *cleaning*
@tanyarlfes Makanya jangan makan produk *boycott*	Makanya jangan makan produk *boycott*
@gravity_away @heyrazni @grok apakah ada signifikansi antara *boycott* produk israel dan melejitnya bisnis lokal indonesia sebagai subtitusi?	apakah ada signifikansi antara *boycott* produk israel dan melejitnya bisnis lokal indonesia sebagai subtitusi
@kamilagimaem_ @BELIFTLAB@ENHYPEN @ENHYPEN_members yg *boycott* produk nya ya kak bukan artisnya *please boycott this collaboration*!!	yg *boycott* produk nya ya kak bukan artisnya *please boycott this collaboration*

2. Case Folding

Tahap selanjutnya *Case folding*, yaitu proses mengubah karakter alfabet yang telah melewati tahap pembersihan menjadi huruf kecil (*lowercase*)(Gifari et al., 2022). Tahapan ini penting dilakukan untuk menghindari perbedaan arti antara kata yang sama namun memiliki penulisan huruf besar dan kecil, sehingga dapat menyederhanakan proses analisis teks dan meningkatkan konsistensi data, seperti yang diperlihatkan pada Tabel 2.

Tabel 2. Hasil *case folding*

Sebelum *case folding*	Setelah *case folding*
Para orang bodoh yang masih menggunakan X *Google* WA dll yang dari Amerika tapi dengan lantang mulut munafik mereka mengucapkan *boycott* produk yang berafiliasi dengan Israel	para orang bodoh yang masih menggunakan x *google* wa dll yang dari amerika tapi dengan lantang mulut munafik mereka mengucapkan *boycott* produk yang berafiliasi dengan israel
Minimal kalau gak mau atau gak bisa ngomong diem gitu dan jangan posting produk *boycott* atau kerjasama dengan zionis kek itu benerbener bare minimum sebagai selebriti yg punya pengaruh besar ke fansnya	minimal kalau gak mau atau gak bisa ngomong diem gitu dan jangan posting produk *boycott* atau kerjasama dengan zionis kek itu benerbener bare minimum sebagai selebriti yg punya pengaruh besar ke fansnya
Lagian tau drmana mereka pake produk *boycott* CocaCola gwe sebagai pengganti beli BIG Cola Gwe pake Mama lemon lebih murah lagi	lagian tau drmana mereka pake produk boycott cocacola gwe sebagai pengganti beli big cola gwe pake mama lemon lebih murah lagi

3. Tokenization

Proses *tokenizing* yang ditunjukkan oleh Tabel 3 merupakan proses memisahkan kalimat menjadi kata-kata individu berdasarkan spasi, dilakukan setelah teks melalui tahap *case folding*(Permata Aulia et al., 2021). Tahap ini penting untuk mempermudah analisis teks karena setiap kata yang telah dipisahkan dapat diproses secara terpisah, seperti dalam proses penghapusan kata-kata tidak penting (*stopword removal*) atau pemberian bobot pada kata. *Tokenizing* juga menjadi dasar bagi proses-proses selanjutnya dalam *text mining* dan analisis sentimen.

Tabel 3. Hasil *tokenization*

Sebelum *tokenization*	Setelah *tokenization*
keepboycotting guys ttp *boycott* ya apapun itu gausah beli produk zionis stop funding ke mereka	*keepboycotting*,guys,ttp,*boycott*,ya,apapun,itu,gausah,beli,produk,zionis,stop,funding,ke,mereka
yg *boycott* produk nya ya kak bukan artisnya *please boycott this collaboration*	yg,*boycott*,produk,nya,ya,kak,bukan,artisnya,*please*,*boycott*,*this*,*collaboration*
makanya jangan makan produk *boycott*	makanya,jangan,makan,produk,*boycott*

4. Stopword Removal

Stopword removal digunakan untuk menghapus kata-kata yang dianggap tidak memiliki makna signifikan, seperti “dan”, “yang”, atau

"atau"(Merdiansah & Ali Ridha, 2024). Proses filtering pada Tabel 4 ini juga membantu mengurangi dimensi data teks yang dianalisis, sehingga mempermudah dan mempercepat proses pengolahan data selanjutnya(Permata Aulia et al., 2021).

Tabel 4. Hasil *stopword removal*

Sebelum *stopword removal*	**Setelah *stopword removal***
ga,cuma,*boycott*,yg,bisa,berhentiin,genosida,cil,tapi,itu,mungkin,itu,salah,satunya,selagi,kita,*boycott*,ada,sdm,yg,terdampak,alhamdulillah,ada,produk,lokal,yg,menyamai,insyaallah,menggantikan,lapangan,kerjanya	ga,*boycott*,yg,berhentiin,genosida,cil,salah,satunya,selagi,*boycott*,sdm,yg,terdampak,alhamdulillah,produk,lokal,yg,menyamai,insyaallah,menggantikan,lapangan,kerjanya
artinya,selama,produk,cepat,tersedia,dan,harganya,terjangkau,sebagian,besar,konsumen,tidak,peduli,apakah,baju,dan,topi,*boycott*,china,dibuat,di,china,atau,amerika,serikat	produk,cepat,tersedia,harganya,terjangkau,konsumen,peduli,baju,topi,*boycott*,china,china,amerika,serikat
gue,masih,*boycott*,btw,*boycott*,produk,atau,*brand*,yg,terafiliasi,sama,isrewel	gue,*boycott*,btw,*boycott*,produk,*brand*,yg,terafiliasi,isrewel

5. Stemming

Pada tahap *stemming* yaitu mencari kata dasar dari sebuah kata(Arsi & Waluyo, 2021). Seperti kata "keresahan" setelah melalui tahap *stemming* menjadi "resah". Pada tahap yang ditunjukkan pada Tabel 5 ini juga berusaha menghilangkan kata-kata yang salah eja dari istilah tersebut.

Tabel 5. Hasil *stemming*

Sebelum *stemming*	**Setelah *stemming***
addition,keresahan,pribadi,aja,tp,sponsor,nya,produk,*boycott*,yaa,kea,air,mineral,yg,cm,*that,one,brand*,yg,ga,beli,karna,panit,nya,ga,bolehin,bawa,air,samsek	*addition*,resah,pribadi,aja,tp,sponsor,nya,produk,*boycott*,yaa,kea,air,mineral,yg,cm,*that,one,brand*,yg,ga,beli,karna,panit,nya,ga,bolehin,bawa,air,samsek
alhamdulillah,kesamaan,sungjin,minyak,telon,nusa,rara,ritual,tidur,ku,rada,ribet,semprot,kasur,biar,wangi,semerbak,ehehe,produk,aman,*boycott*	alhamdulillah,sama,sungjin,minyak,ton,nusa,rara,ritual,tidur,ku,rada,ribet,semprot,kasur,biar,wangi,semerbak,ehehe,produk,aman,*boycott*
more,reason,to,boycott,nih,produk,minuman,gak,enak,gula,bikin,diabet	*more,reason,to,boycott*,nih,produk,minum,gak,enak,gula,bikin,abet

2.3 Pelabelan Data

Setelah data dibersihkan, tahap berikutnya adalah pemberian label sentimen. Komentar-komentar pada tweet dikategorikan ke dalam tiga kelas utama yaitu pro, kontra, dan netral. Pemberian label ini dilakukan secara semi-otomatis dengan validasi manual guna memastikan tingkat akurasi yang tinggi. Tahap ini krusial untuk mempersiapkan data agar siap dianalisis lebih lanjut menggunakan algoritma machine learning(Hariyanto et al., 2020).

Tabel 6. Hasil pelabelan data

['gue', '*boycott*', 'btw', '*boycott*', 'produk', '*brand*', 'yg', 'afiliasi', 'isrewel']	**pro**
['kalo', 'henti', 'genosida', 'maju', 'produk', 'lokal', 'buka', 'lapang', 'kerja', 'orang', 'yg', 'dampak', '*boycott*', 'ngga', 'coba', 'kalo', 'mikir', 'tuh', '*multiplier*', 'effectnya', 'perkara', 'aja']	**kontra**
['iya', 'sih', 'paham', 'khawatir', 'roda', 'ekonomi', '*slow*', '*down*', 'aja', 'klo', 'produk', 'dlm', 'negeri', '*boycott*', 'mdhan', 'produk', 'alternatif', 'ttp', 'gerakin', 'ekonomi', 'ya']	**netral**

2.4 Pembobotan Data

Setelah data melalui proses pelabelan, tahap selanjutnya adalah memberikan bobot atau pembobotan data pada setiap kata dalam data menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini berfungsi untuk mengukur tingkat relevansi atau signifikansi setiap kata dalam analisis, dengan tujuan meningkatkan ketepatan model dalam mengenali sentimen(Ashari et al., 2023). TF-IDF berperan dalam mengurangi pengaruh kata-kata umum yang kurang informatif dan menonjolkan kata-kata spesifik yang relevan, sehingga model dapat lebih fokus pada fitur penting yang mencerminkan masing-masing kelas sentimen.

2.5 Pemisahan Data

Pemisahan data adalah proses membagi dataset menjadi dua bagian atau lebih(Arif Widiasan Subagio et al., 2024), dengan tujuan memisahkan data untuk pelatihan (*training*) dan pengujian (*testing*) model machine learning. Pemisahan data merupakan tahapan penting agar model yang dilatih dapat diuji secara adil dan tidak hanya menghafal data. Dalam penelitian ini, data dibagi dengan proporsi 80% untuk data latih dan 20% untuk data uji, menghasilkan 3352 data latih dan 838 data uji dengan perbandingan 80:20.

2.6 Pemodelan SVM

Data yang telah diberi bobot menggunakan metode TF-IDF digunakan untuk melatih model dengan SVM. SVM adalah metode machine learning baru yang digunakan dalam melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. Secara umum, SVM bekerja dengan pendekatan statistik dalam mengenali pola dari data yang ada. SVM akan mencari support vector terbaik yang bisa memisahkan dua kelompok data berbeda dengan jarak atau margin yang paling lebar. Salah satu keunggulan utama SVM adalah performanya yang cenderung lebih baik dibanding teknik klasifikasi lain, terutama karena pendekatan matematisnya yang kuat dan terstruktur dengan jelas(Tinaliah & Elizabeth, 2022).

2.7 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma SVM dalam mengklasifikasikan sentimen terhadap gerakan boikot di platform X. Untuk itu, model akan diuji melalui empat skenario pengujian yang menggabungkan penggunaan dua jenis kernel, yaitu kernel linier dan kernel Radial Basis Function (RBF), serta metode balancing data menggunakan SMOTE. Keempat skenario tersebut mencakup, penggunaan kernel linier tanpa SMOTE, penggunaan kernel linier dengan SMOTE, penggunaan kernel RBF tanpa SMOTE, dan penggunaan kernel RBF dengan SMOTE. Tujuan dari skenario pengujian ini adalah untuk mengevaluasi sejauh mana metode balancing dan pemilihan kernel mempengaruhi kemampuan model dalam mengenali dan memisahkan kelas sentimen pro, netral, dan kontra secara efektif.

Untuk mengukur performa model pada setiap skenario, digunakan sejumlah *metric* evaluasi yang umum dalam klasifikasi teks, yaitu *accuracy, precision, recall,* dan *F1-score*. Keempat *metric* ini digunakan untuk menilai tingkat ketepatan dan kemampuan model dalam mengenali tiap kelas sentimen, khususnya pada data yang tidak seimbang. Selain itu, digunakan juga *confusion matrix* untuk melihat distribusi prediksi benar dan salah antar kelas. Untuk klasifikasi multi-kelas, ditambahkan *metric macro average* dan *weighted average* yang digunakan untuk mengevaluasi kinerja model secara merata dan proporsional. Kombinasi skenario dan *metric* ini diharapkan dapat memberikan gambaran menyeluruh dan objektif terhadap kemampuan model dalam mengklasifikasikan sentimen pro, netral, dan kontra.

3. Hasil dan Pembahasan

3.1 TF-IDF

Matriks TF-IDF yang disajikan dalam laporan ini menggambarkan representasi numerik dari 4190 dokumen yang dianalisis, dengan 7895 kata unik yang berfungsi sebagai fitur yang bisa dilihat pada Gambar 3. Hasil pembobotan menggunakan metode TF-IDF menunjukkan bahwa kata-kata seperti "boycott", "produk", dan "Israel" memiliki bobot yang tinggi di sebagian besar dokumen. Hal ini menandakan bahwa ketiga kata tersebut merupakan isu utama yang dominan dalam percakapan warganet terkait gerakan boikot di platform X. Selain itu, kata-kata seperti "lokal" dan "alternatif" juga muncul dengan bobot yang cukup tinggi, menggambarkan arah opini publik yang tidak hanya menolak produk tertentu, tetapi juga mendukung produk lokal sebagai pengganti. Dengan demikian, TF-IDF menjadi dasar penting dalam mengubah data teks menjadi bentuk numerik yang dapat diproses oleh model analitik lanjutan.

Matriks TF-IDF:

	aaaa	aaaaa	aaakjksjdisjidjsjf	aamiin	aammiin	ab	abah	abai	abal	\
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
2	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
4	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...		...	...	...	...	...	...	...
4185	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
4186	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
4187	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
4188	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
4189	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0

	abang	...	zone	zonis	zonist	zonk	zonst	zranya	zs	zus	zwitsal	\
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	
...	...	...	...	...	...	...	...	...	...	...	...	
4185	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4186	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4187	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4188	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4189	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

	zwitsl
0	0.0
1	0.0
2	0.0
3	0.0
4	0.0
...	...
4185	0.0
4186	0.0
4187	0.0
4188	0.0
4189	0.0

[4190 rows x 7895 columns]

Gambar 3. Hasil pembobotan data

Nilai TF-IDF yang tinggi menunjukkan bahwa suatu kata sangat relevan untuk dokumen tertentu, sedangkan nilai mendekati nol mengindikasikan bahwa kata tersebut kurang relevan atau tidak muncul dalam dokumen tersebut. Karena jumlah kata yang unik dalam koleksi cukup banyak, matriks ini bersifat sparse, artinya banyak nilai yang berisi nol, karena tidak semua kata muncul di setiap dokumen.

```
Fitur (kata-kata):
['aaaa' 'aaaaa' 'aaakjksjdisjidjsjf' ... 'zus' 'zwitsal' 'zwitsl']
```

Gambar 4. Daftar kata-kata pada pembobotan data

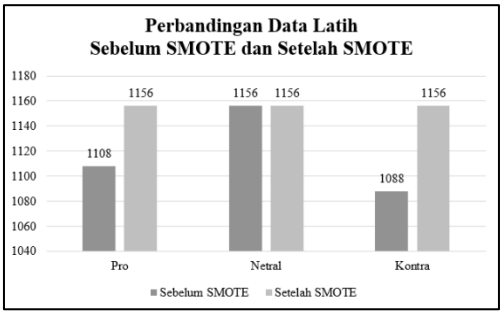
Daftar kata-kata yang ada di dalam fitur menunjukkan kata-kata unik yang telah dipilih dari seluruh kumpulan dokumen yang dapat dilihat pada Gambar 4, yang digunakan untuk menyusun kolom dalam matriks. Misalnya, jika kata 'aamiin' muncul dalam satu dokumen, kolomnya akan menunjukkan nilai TF-IDF yang lebih tinggi dibandingkan dengan kata-kata yang lebih jarang muncul. Dengan cara ini, TF-IDF membantu mengidentifikasi kata-kata yang lebih relevan atau penting di dalam setiap dokumen dalam konteks seluruh koleksi dokumen yang ada.

3.2 Perbandingan Akurasi

Setelah melalui tahap pra-pemrosesan, penelitian ini memperoleh 3352 data latih yang terdiri atas 1108 data bersentimen pro, 1156 data netral, dan 1088 data kontra, sebagaimana ditunjukkan pada Gambar 5. Ketidakseimbangan distribusi jumlah data pada masing-masing kelas sentimen menyebabkan algoritma SVM cenderung lebih optimal dalam mempelajari pola dari kelas mayoritas, yaitu sentimen netral. Hal ini berdampak pada performa model yang tinggi pada kelas tersebut, namun kurang akurat dalam mengklasifikasikan sentimen pro dan kontra.

Untuk mengatasi ketimpangan ini, penelitian menerapkan teknik optimasi SMOTE, yang berfungsi menyeimbangkan jumlah data latih dengan menambahkan data sintetik pada kelas

minoritas(Saputra, 2025). Penerapan SMOTE berhasil menyamakan jumlah data pada ketiga kelas sentimen menjadi masing-masing 1156 data, sehingga model dapat belajar dari distribusi yang lebih seimbang. Hasil distribusi data latih dengan kernel linier dan RBF sebelum dan sesudah SMOTE ditampilkan pada Gambar 5.



Gambar 5. Perbandingan data latih sebelum SMOTE dan sesudah SMOTE

3.3 Tahap Pengujian

Penelitian ini mengevaluasi performa terbaik dari setiap model algoritma dan membagi data menjadi dua bagian, yaitu 80% (3352 data) sebagai data pelatihan dan 20% (838 data) sebagai data pengujian yang diterapkan pada algoritma SVM. Seluruh data telah melalui proses optimasi menggunakan teknik SMOTE, sehingga distribusi sentimen menjadi lebih seimbang dan model dapat belajar secara lebih adil dari tiap kelas sentimen.

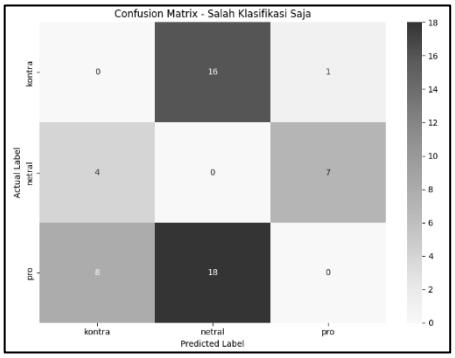
Tahap selanjutnya adalah melakukan interpretasi terhadap hasil klasifikasi yang diperoleh dari algoritma SVM. Pada tahap ini, dilakukan analisis terhadap hasil klasifikasi untuk memahami pola dan tren yang berhasil dikenali oleh model, serta untuk menilai akurasi dan kinerja model dalam memprediksi kategori atau label dari data yang tersedia(Afandi et al., 2024). Analisis difokuskan pada penerapan algoritma SVM dengan dua jenis kernel, yaitu kernel linier dan kernel non-linier (RBF), baik sebelum maupun sesudah diterapkannya teknik SMOTE untuk penyeimbangan data.

3.3.1. Kinerja Model SVM dengan Kernel Linier

Pada kernel linier, penerapan SMOTE memberikan dampak yang cukup signifikan terhadap peningkatan kinerja model klasifikasi. Gambar 7 menunjukkan hasil klasifikasi model SVM dengan kernel linier sebelum dilakukan penyeimbangan data menggunakan teknik SMOTE. Model menghasilkan akurasi sebesar 92%, dengan *precision* tertinggi pada kelas pro 0.98, diikuti oleh kelas kontra 0.93 dan kelas netral 0.88. Dari sisi *recall*, kelas netral menunjukkan performa terbaik dengan nilai 0.96, disusul oleh kelas kontra 0.92, sedangkan kelas pro menunjukkan nilai yang relatif lebih rendah yaitu 0.89, yang menandakan bahwa model masih kesulitan mengenali kelas pro.

Hal tersebut diperjelas melalui Gambar 6 yang menunjukkan confusion matrix salah klasifikasi. Dari total 54 data yang salah diklasifikasikan, terlihat bahwa model memiliki kecenderungan kuat untuk mengarahkan prediksi ke kelas netral, terutama dari kelas pro dan kontra. Sebanyak 18 data kelas pro diklasifikasikan sebagai netral, dan 8 data pro lainnya salah dikenali sebagai kontra. Selain itu, terdapat 16 data kontra yang juga diklasifikasikan sebagai netral, serta 7 data netral yang justru dikenali sebagai pro.

Contoh nyata dari kesalahan klasifikasi ini diperlihatkan oleh data dengan indeks 3632 yang berisi kalimat *"sama sekali bukan membelamembenarkan idol kpop..."*, yang semestinya berlabel pro, namun diklasifikasikan sebagai netral. Begitu pula data 3449, yang menyatakan *"jangan generalisir ya kak gue fandom ungu tapi..."*, juga salah diklasifikasikan sebagai netral. Analisis terhadap konten menunjukkan bahwa kalimat dalam kelas pro sering menggunakan ekspresi dukungan secara halus atau implisit, yang membuat model kesulitan membedakan dari pernyataan netral.



Gambar 6. Confusion matrix data salah klasifikasi

Nilai *F1-score* tertinggi dimiliki oleh kelas pro yaitu 0.93, sementara kelas kontra dan netral sama-sama memperoleh nilai 0.92. Nilai *macro average F1-score* adalah 0.93, dan *weighted average*-nya adalah 0.92. Meskipun akurasi tinggi, *confusion matrix* mengindikasikan adanya kecenderungan model untuk lebih mengklasifikasikan data ke dalam kelas mayoritas, khususnya netral. Hal ini merupakan indikasi dari bias akibat ketidakseimbangan data, yang dapat mempengaruhi generalisasi model terhadap kelas minoritas.

```
Akurasi: 0.9248210023866349

Classification Report:
              precision    recall  f1-score   support

      kontra       0.93      0.92      0.92       300
      netral       0.88      0.96      0.92       296
         pro       0.98      0.89      0.93       242

    accuracy                           0.92       838
   macro avg       0.93      0.92      0.93       838
weighted avg       0.93      0.92      0.92       838


Confusion Matrix:
 [[275  25   0]
 [  7 284   5]
 [ 14  12 216]]
```

Gambar 7. Hasil svm (kernel linier) sebelum smote

Distribusi label menjadi lebih seimbang dan berdampak pada peningkatan performa model secara keseluruhan setelah diterapkan teknik SMOTE, seperti yang ditampilkan pada Gambar 8. Setelah

dilakukan *oversampling* terhadap kelas minoritas, performa model mengalami peningkatan dengan akurasi mencapai 94%. Terjadi perbaikan yang cukup signifikan pada seluruh metrik evaluasi. Nilai *Precision* meningkat hampir merata pada ketiga kelas, dengan nilai tertinggi pada kelas pro 0.97, lalu kontra 0.96, dan netral 0.89. Dari sisi nilai *recall*, kelas netral tetap mendominasi yaitu 0.96, namun kelas pro dan kontra juga menunjukkan peningkatan menjadi masing-masing 0.91 dan 0.94. Hal ini menunjukkan bahwa model kini dapat mengenali *instance* dari setiap kelas dengan lebih seimbang. Nilai *F1-score* tertinggi dimiliki oleh kelas kontra 0.95, diikuti oleh pro kelas 0.94 dan kelas netral 0.93. Baik macro average maupun *weighted average F1-score* meningkat menjadi 0.94, yang mencerminkan bahwa model telah menjadi lebih adil dan konsisten dalam mengklasifikasikan ketiga kelas secara proporsional. *Confusion matrix* juga memperlihatkan distribusi prediksi yang lebih seimbang, menandakan bahwa penggunaan SMOTE berhasil mengurangi bias terhadap kelas mayoritas dan meningkatkan kinerja keseluruhan model.

```
Akurasi: 0.9355608591885441

Classification Report:
              precision    recall  f1-score   support

      kontra       0.96      0.94      0.95       272
      netral       0.89      0.96      0.93       289
         pro       0.97      0.91      0.94       277

    accuracy                           0.94       838
   macro avg       0.94      0.94      0.94       838
weighted avg       0.94      0.94      0.94       838


Confusion Matrix:
 [[255  16   1]
 [  4 278   7]
 [  8  18 251]]
```

Gambar 8. Hasil svm (kernel linier) setelah smote

3.3.2. Kinerja Model SVM dengan Kernel RBF

Setelah dilakukan evaluasi terhadap kernel linier, analisis dilanjutkan dengan kernel RBF untuk membandingkan performa antar jenis kernel. Gambar 9 menampilkan hasil klasifikasi SVM dengan kernel RBF sebelum dilakukan SMOTE. Model mencapai akurasi sebesar 92%, dengan *precision* tertinggi pada kelas pro 1.00, yang menunjukkan bahwa semua prediksi kelas pro adalah benar, namun *recall*-nya lebih rendah 0.86, mengindikasikan bahwa masih banyak data kelas pro yang gagal dikenali. *Precision* pada kelas kontra juga tinggi 0.97, disusul oleh kelas netral yang paling rendah 0.84. Dari sisi *recall*, kelas netral menunjukkan performa terbaik dengan nilai 0.99, yang berarti hampir semua *instance* kelas netral berhasil dikenali oleh model. Kelas kontra memiliki *recall* 0.92, dan kelas pro memiliki nilai *recall* terendah yaitu 0.86. Nilai *F1-score* model cukup merata dengan nilai 0.94 untuk kelas kontra, 0.91 untuk kelas netral, dan 0.92 untuk kelas pro. Nilai macro average *F1-score* mencapai 0.93, menunjukkan performa umum model terhadap semua kelas cukup baik, meskipun ada indikasi ketimpangan dalam pengenalan kelas pro. Hal ini diperkuat dengan *confusion matrix*, di mana banyak *instance* pro 32 data justru diklasifikasikan sebagai netral, menunjukkan potensi bias akibat ketidakseimbangan jumlah data antar kelas.

```
Akurasi: 0.9248210023866349

Classification Report:
              precision    recall  f1-score   support

      kontra       0.97      0.92      0.94       272
      netral       0.84      0.99      0.91       289
         pro       1.00      0.86      0.92       277

    accuracy                           0.92       838
   macro avg       0.94      0.92      0.93       838
weighted avg       0.93      0.92      0.93       838


Confusion Matrix:
 [[249  23   0]
 [  1 287   1]
 [  6  32 239]]
```

Gambar 9. Hasil svm (rbf) sebelum smote

Gambar 10 menampilkan hasil klasifikasi setelah diterapkannya SMOTE untuk menyeimbangkan jumlah data antar kelas. Menariknya, hasil evaluasi tidak menunjukkan perubahan signifikan dibanding sebelum SMOTE, baik dari segi akurasi, *precision*, *recall*, maupun *F1-score.* Akurasi tetap berada di angka 92%, precision dan *recall* tiap kelas pun sama persis dengan nilai sebelum SMOTE. Nilai *macro average* dan *weighted average F1-score* juga tetap sebesar 0.93, dengan distribusi performa model yang serupa. Confusion matrix pun identik, menunjukkan bahwa dalam kasus ini, SMOTE tidak memberikan dampak signifikan terhadap peningkatan performa model. Hal ini bisa terjadi karena distribusi awal sudah cukup seimbang atau karena fitur-fitur hasil vektorisasi sudah cukup representatif tanpa perlu oversampling tambahan. Namun demikian, penerapan SMOTE tetap memberikan jaminan bahwa model tidak bias terhadap kelas minoritas dalam skenario data nyata yang lebih timpang di masa depan.

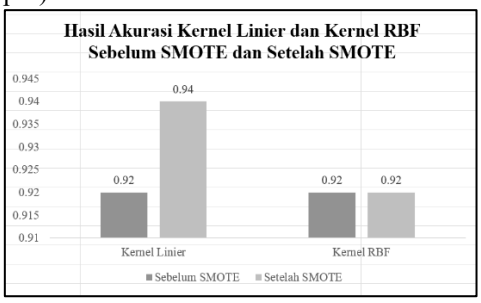
```
Akurasi: 0.9248210023866349

Classification Report:
              precision    recall  f1-score   support

      kontra       0.97      0.92      0.94       272
      netral       0.84      0.99      0.91       289
         pro       1.00      0.86      0.92       277

    accuracy                           0.92       838
   macro avg       0.94      0.92      0.93       838
weighted avg       0.93      0.92      0.93       838


Confusion Matrix:
 [[249  23   0]
 [  1 287   1]
 [  6  32 239]]
```

Gambar 10. Hasil svm (rbf) setelah smote

3.3.3. Perbandingan Kinerja Kernel Linier dengan RBF

Berdasarkan Gambar 11 performa model pada keempat skenario pengujian, didapatkan hasil model **SVM kernel linier** yang menerapkan SMOTE memberikan performa terbaik dibandingkan konfigurasi lainnya. Model ini mencatatkan akurasi sebesar 0.94, dengan nilai *macro average F1-score*

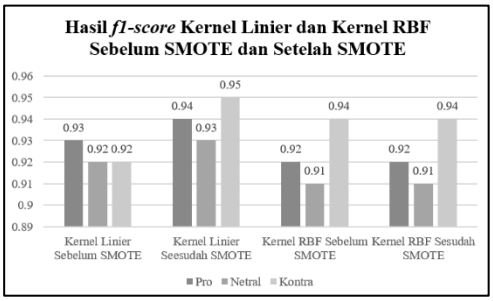
sebesar 0.94 dan *macro average recall* sebesar 0.94, yang mengindikasikan keseimbangan klasifikasi yang lebih baik di antara seluruh kelas (kontra, netral, dan pro).



Gambar 11. Perbandingan hasil akurasi kernel linier dan RBF

Gambar 12 menunjukkan perbandingan nilai *f1-score* dari algoritma SVM dengan dua jenis kernel, yaitu kernel linier dan kernel RBF, sebelum dan sesudah diterapkannya teknik SMOTE. Sebelum SMOTE diterapkan, kernel linier menunjukkan nilai *f1-score* yang relatif seimbang pada ketiga kelas, yakni pro 0.93, netral 0.92, dan kontra 0.92. Namun, setelah penerapan SMOTE, terjadi peningkatan yang cukup signifikan, terutama pada kelas kontra yang naik menjadi 0.95, sementara kelas pro dan netral meningkat menjadi masing-masing 0.94 dan 0.93. Hal ini menunjukkan bahwa kernel linier mampu memanfaatkan data sintetis hasil SMOTE untuk meningkatkan generalisasi terhadap kelas minoritas.

Sementara itu, pada kernel RBF, nilai *f1-score* untuk kelas pro dan netral masing-masing adalah 0.92 dan 0.91 baik sebelum maupun sesudah SMOTE, sementara kelas kontra tetap stabil pada angka 0.94. Tidak adanya peningkatan signifikan pada kernel RBF setelah penerapan SMOTE yang mengindikasikan bahwa kernel ini tetap bekerja baik meskipun data tidak seimbang. Secara keseluruhan, hasil ini menunjukkan bahwa SMOTE lebih efektif meningkatkan performa kernel linier dibandingkan kernel RBF.



Gambar 12. Perbandingan hasil *f1-score* kernel linier dan RBF

Model SVM dengan kernel linier menunjukkan performa yang sangat baik dalam klasifikasi, tidak hanya dari segi akurasi yang tinggi, tetapi juga dari keseimbangan hasil prediksi antar kelas. Keunggulan kernel linier dalam penelitian ini berkaitan erat dengan karakteristik data yang secara alami memiliki pola distribusi yang cenderung linier, sehingga pemisahan antar kelas dapat dicapai tanpa memerlukan transformasi non-linier yang kompleks. Berbeda dengan kernel RBF yang dirancang untuk menangani data non-linier, model linier justru bekerja lebih optimal ketika struktur data relatif sederhana. Pada konteks ini, penerapan RBF linier berpotensi menyebabkan overfitting, terutama jika kompleksitas pemetaan ruang fitur tidak sesuai dengan struktur data yang sebenarnya. Hal ini menjelaskan mengapa kernel linier dapat mencapai akurasi yang lebih tinggi dibanding kernel RBF.

Penerapan SMOTE juga berkontribusi terhadap peningkatan performa model secara keseluruhan. Sebelum SMOTE, ketidakseimbangan distribusi kelas menyebabkan model lebih condong mengenali kelas mayoritas, sehingga performa pada kelas minoritas relatif rendah. Setelah dilakukan SMOTE, distribusi data antar kelas menjadi seimbang dan memungkinkan model khususnya kernel linier untuk mempelajari representasi dari setiap kelas secara lebih proporsional. Dengan demikian, kombinasi antara kernel linier dan SMOTE terbukti mampu menghasilkan model klasifikasi yang lebih stabil, akurat, dan generalis, khususnya dalam menghadapi tantangan data yang tidak seimbang. Evaluasi performa secara menyeluruh memperkuat bahwa kernel linier merupakan pilihan yang sangat tepat dalam penelitian ini.

Namun demikian, penelitian ini memiliki beberapa keterbatasan yang perlu dicermati. Model yang digunakan terbatas pada algoritma SVM dengan representasi fitur berbasis TF-IDF, yang belum mampu sepenuhnya menangkap makna kontekstual atau nuansa semantik dalam teks. Hal ini dapat mempengaruhi performa model, khususnya dalam mengklasifikasikan kelas pro yang cenderung menggunakan bahasa tidak eksplisit atau ambigu. Selain itu, ruang lingkup data yang hanya bersumber dari platform X dalam kurun waktu tertentu membatasi generalisasi hasil terhadap konteks media sosial yang lebih luas. Ke depannya, penelitian serupa dapat dikembangkan menggunakan model berbasis deep learning seperti BERT atau arsitektur transformer lainnya yang lebih sensitif terhadap konteks kalimat. Penambahan analisis temporal, perluasan sumber data ke platform lain, serta penerapan teknik multi-label classification juga dapat menjadi arah pengembangan yang potensial dalam memahami kompleksitas opini publik terhadap isu sosial-politik di ranah digital.

4. Kesimpulan

Hasil penelitian menunjukkan bahwa metode SVM mampu mengklasifikasikan sentimen masyarakat terhadap gerakan boycott di platform X ke dalam tiga kelas utama, yaitu pro, netral, dan kontra. Dengan total 4190 data dengan 838 data uji, penerapan teknik SMOTE berhasil mengatasi ketidakseimbangan data antar kelas, sehingga menghasilkan distribusi data yang lebih merata

sebanyak 1156 data. Hasil pengujian menunjukkan bahwa model SVM mampu mengklasifikasikan data dengan tingkat akurasi tertinggi sebesar 94% pada kombinasi kernel linier dengan SMOTE, yang juga menunjukkan performa paling stabil di seluruh kelas. Pada skenario lain, akurasi model tercatat sebesar 92% pada kernel linier tanpa SMOTE, 92% pada kernel RBF tanpa SMOTE, dan 92% pada kernel RBF dengan SMOTE. Perbandingan ini memperlihatkan bahwa penggunaan SMOTE meningkatkan performa model secara signifikan pada kernel linier, namun tidak memberikan dampak serupa pada kernel RBF. Peningkatan performa pada kernel linier terjadi karena model ini sangat bergantung pada distribusi data yang seimbang untuk membentuk garis pemisah yang optimal. Ketika data tidak seimbang, model cenderung bias ke kelas mayoritas. SMOTE membantu dengan menambahkan data sintetis pada kelas minoritas, sehingga batas antar kelas bisa dipelajari lebih proporsional. Sebaliknya, kernel RBF yang fleksibel mampu membentuk batas non-linier dan tetap bekerja baik meski data tidak seimbang, sehingga dampak SMOTE tidak terlalu signifikan dibandingkan pada kernel linier. Dengan demikian, kombinasi SVM dengan kernel linier dan SMOTE merupakan pilihan terbaik dan terbukti lebih sesuai dalam konteks penelitian ini, terutama dalam menangani distribusi data yang tidak seimbang dan menangkap pola sentimen yang ada secara lebih akurat dan generalis.

Daftar Pustaka:

Afandi, R. B., Nurdiansyah, T. F., Ramadhani, A. N., & Sari, A. P. (2024). Implementasi Support Vector Machine untuk Analisis Sentimen Aplikasi "Mpstore - Super App UMKM." *Jurnal Informatika Polinema*, *10*(4), 565–570. https://doi.org/10.33795/jip.v10i4.5427

Alifa, C. F., & Alita, D. (2024). Analisis Opini Publik Tentang Boikot Produk Pro-Israel di Twitter Berbahasa Indonesia Menggunakan Metode SVM. *Jurnal Informatika: Jurnal Pengembangan IT*, *9*(2), 112–120. https://doi.org/10.30591/jpit.v9i2.6559

Arif Widiasan Subagio, Anggraini Puspita Sari, & Andreas Nugroho Sihananto. (2024). Klasifikasi Lexicon-Based Sentiment Analysis Tragedi Kanjuruhan pada Twitter Menggunakan Algoritma Convolutional Neural Network. *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, *4*(1), 166–177. https://doi.org/10.55606/juisik.v4i1.759

Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(1), 147. https://doi.org/10.25126/jtiik.0813944

Ashari, S. A., Saputra, M. W. A., Larosa, E., & Rijal, B. S. (2023). Analisis Sentimen pada Aplikasi Translate Google Menggunakan Metode SVM (Studi Kasus: Komentar Pada Playstore). *Jurnal Teknik*, *21*(2), 168–182. https://doi.org/10.37031/jt.v21i2.412

Aziz, A. (2022). Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *6*(1), 115. https://doi.org/http://dx.doi.org/10.30645/j-sakti.v6i1.430

Fitri, F. A., & Ulmuftia, N. (2025). *Peran Media Sosial dalam Mendorong Perilaku Konsumen terhadap Boikot Produk Israel*. *5*(2), 72–81. https://doi.org/https://doi.org/10.31764/jseit.v5i2.30659

Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine. *Journal of Information Technology*, *2*(1), 36–40. https://doi.org/https://doi.org/10.46229/jifotech.v2i1.330

Hariyanto, M., Kholiq, M., Yani, A., & Narti. (2020). Inti nusa mandiri. *Inti Nusa Mandiri*, *14*(2), 133–138. https://doi.org/https://doi.org/10.33480/inti.v19i2.6297

Hendrastuty, N., Rahman Isnain, A., & Yanti Rahmadhani, A. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT*, *6*(3), 150–155. https://doi.org/https://doi.org/10.30591/jpit.v6i3.2870

Idris Siregar, Dinar Maliki, & Fakhrurrozi Nasution. (2024). Jihad Ekonomi Dalam Perspektif Hadis: Tinjauan Terhadap Gerakan Boikot Produk Israel Sebagai Dukungan Terhadap Palestina. *MUQADDIMAH: Jurnal Ekonomi, Manajemen, Akuntansi Dan Bisnis*, *2*(3), 181–192. https://doi.org/10.59246/muqaddimah.v2i3.938

Merdiansah, R., & Ali Ridha, A. (2024). Sentiment Analysis of Indonesian X Users Regarding Electric Vehicles Using IndoBERT. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI*, *7*(1), 221–228. https://doi.org/https://doi.org/10.55338/jikomsi.v7i1.2895

Munandar, A., Yaasin, M. S., & Firdaus, R. A. (2023). Analisis Sentimen Netizen Indonesia Mengenai Boikot Produk. *Journal of Islamic Banking and Economics*, *3*(1), 23–40. https://doi.org/https://doi.org/10.15408/thd.v3i1.37381

Permata Aulia, T. M., Arifin, N., & Mayasari, R.

(2021). Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinisasi Covid-19. *SINTECH (Science and Information Technology) Journal*, *4*(2), 139–145. https://doi.org/10.31598/sintechjournal.v4i2.762

PS, A., Nugroho Sihananto, A., & Arman Prasetya, D. (2022). Implementasi Metode K-NN dalam Klasterisasi Kasus Kesehatan Jantung. *ALINIER: Journal of Artificial Intelligence & Applications*, *3*(2), 18–21. https://doi.org/10.36040/alinier.v3i2.5761

Ray, S., Das, J., Pande, R., & Nithya, A. (2024). *Swati Ray 1 , Joyati Das 2* , Ranjana Pande 3 , and A. Nithya 2*. *3*, 195–222. https://doi.org/10.1201/9781032622408-13

Saputra, M. R. (2025). *Analisis sentimen twitter terhadap konflik di papua menggunakan perbandingan naive bayes dan svm*. *10*(2), 1197–1208. https://doi.org/https://doi.org/10.29100/jipi.v10i2.6180

Susilawati, A. T., Lestari, N. A., & Nina, P. A. (2024). Analisis Sentimen Publik Pada Twitter Terhadap Boikot Produk Israel Menggunakan Metode Naïve Bayes. *Jurnal Ilmiah Mahasiswa*, *2*(1), 26–35. https://doi.org/10.59603/niantanasikka.v2i1.240

Tantika, R. S., & Kudus, A. (2022). Penggunaan Metode Support Vector Machine Klasifikasi Multiclass pada Data Pasien Penyakit Tiroid. *Bandung Conference Series: Statistics*, *2*(2), 159–166. https://doi.org/10.29313/bcss.v2i2.3590

Tarigan, V. T., & Yusupa, A. (2024). Perbandingan Algoritma Maching Learning dalam Analisis Sentimen Mobil Listrik di Indonesia pada Media Sosial Twitter/X. *Jurnal Informatika Polinema*, *10*(4), 479–490. https://doi.org/10.33795/jip.v10i4.5130

Tinaliah, T., & Elizabeth, T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *9*(4), 3436–3442. https://doi.org/10.35957/jatisi.v9i4.3586