

Klasifikasi Sentimen Pengguna X Terhadap Pemboikotan Produk Pro Israel Menggunakan Algoritma Machine Learning

Pingki Muliya Susanti*, M Afdal, Inggih Permana, Arif Marsal

Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia

Email: 12150320272@students.uin-suska.ac.id, m.afdal@uin-suska.ac.id, inggihpermana@uin-suska.ac.id, arif.marsal@uin-suska.ac.id

Email Penulis Korespondensi: 12150320272@students.uin-suska.ac.id

Submitted: 26/12/2024; Accepted: 26/02/2025; Published: 01/03/2025

Abstrak—Gerakan boikot produk Pro Israel muncul akibat konflik berkepanjangan antara Israel dan Palestina. Gerakan boikot berdampak pada penurunan penjualan yang berimbas pada kesejahteraan karyawan seperti berkurangnya bonus, pemotongan gaji, dan pemutusan hubungan kerja (PHK). Tindakan ini memicu beragam tanggapan dari masyarakat di platform X. Penelitian ini bertujuan untuk mengklasifikasikan sentimen pengguna X terhadap pemboikotan produk pro-Israel dengan membandingkan kinerja algoritma Machine Learning yaitu Support Vector Machine dan Random Forest. Untuk mengatasi ketidakseimbangan kelas dalam dataset, penelitian ini menggunakan metode synthetic minority over-sampling technique (SMOTE). Data yang digunakan berjumlah 2.275, diperoleh melalui teknik crawling pada platform X. Hasil penelitian menunjukkan bahwa mayoritas pengguna X di Indonesia mendukung gerakan boikot dengan sentimen positif sebesar 58%. Algoritma SVM dengan SMOTE memberikan kinerja terbaik dalam mengklasifikasikan sentimen dengan akurasi mencapai 90,54%, sedangkan Random Forest hanya memperoleh akurasi 83,1%. Penelitian ini memberikan pemahaman tentang perspektif masyarakat Indonesia terhadap aksi boikot produk pro-Israel.

Kata Kunci: Boikot Produk Israel; Klasifikasi Sentimen; Machine Learning; Random Forest; Support Vector Machine

Abstract—The campaign to boycott pro-Israel goods emerged as a result of the enduring conflict between Israel and Palestine. This boycott initiative led to a decline in sales, which adversely impacted the livelihoods of employees, manifesting in diminished bonuses, salary reductions, and job terminations. Such actions elicited a variety of reactions from the public on platform X. This study seeks to categorize the sentiments of X users regarding the boycott of pro-Israel products by comparing the efficacy of Machine Learning algorithms, namely Support Vector Machine and Random Forest. To address the class imbalance within the dataset, this research employs the synthetic minority over-sampling technique (SMOTE). The dataset comprised 2,275 entries, gathered through web scraping methods on the X platform. The findings indicate that a majority of X users in Indonesia endorse the boycott movement, exhibiting a positive sentiment of 58%. The SVM algorithm, when combined with SMOTE, demonstrated the highest performance in sentiment classification, achieving an accuracy of 90.54%, whereas Random Forest attained an accuracy of only 83.1%. This research offers insights into the views of the Indonesian populace regarding the boycott of pro-Israel products.

Keywords: Boycott Israeli Products; Sentiment Classification; Machine Learning; Random Forest; Support Vector Machine

1. PENDAHULUAN

Konflik berkepanjangan antara Israel dan Palestina telah berlangsung lama sejak tahun 1917 hingga saat ini [1]. Konflik ini melibatkan perebutan wilayah, serta tindakan kekerasan yang merampas hak-hak dasar manusia. Berbagai aksi, pernyataan kecaman, dan kritik ditujukan kepada Israel oleh negara-negara Islam. Pada November 2023, Organisasi Konferensi Islam (OKI) mengecam Israel dan mendesak anggota OKI untuk mengambil tindakan hukum terhadap Israel atas kejahatan perang [2]. Meski upaya untuk mencapai solusi perdamaian masih menemui jalan buntu, berbagai kecaman dan kritikan terus mengalir terhadap Israel termasuk upaya gerakan boikot.

Gerakan boikot terhadap produk Israel maupun produk yang terafiliasi Israel merupakan salah satu bentuk perlawanan. Gerakan ini menyerukan masyarakat untuk tidak menggunakan produk-produk yang mendukung Israel, dengan tujuan untuk memutus hubungan ekonomi, sosial, atau politik dengan negara tersebut [3]. Aksi boikot ini dapat menyebabkan penurunan penjualan produk, yang berdampak negatif pada kesejahteraan karyawan seperti berkurangnya bonus, pemotongan gaji, hingga PHK karyawan. Produk-produk yang sering kali menjadi sasaran dalam gerakan boikot yaitu KFC, Burger King, Pizza Hut, Starbucks, McDonalds, Unilever dan *brand fashion intern* lainnya [4] [5]. Alasan utama di balik aksi boikot ini adalah dugaan bahwa sebagian keuntungan penjualan produk pro-Israel digunakan untuk mendanai kegiatan militer dan pembangunan infrastruktur, serta pengembangan di Israel. [5].

Majelis Ulama Indonesia (MUI) mengeluarkan Fatwa Nomor 83 Tahun 2023 sebagai respon terhadap aksi boikot tersebut. MUI menyatakan bahwa mendukung agresi Israel ke Palestina atau pihak yang mendukung Israel secara langsung maupun tidak langsung adalah haram. MUI juga menghimbau umat Islam untuk menghindari penggunaan barang-barang yang terkait dengan Israel [6]. Menurut Fatwa Rabithah Ulama Palestina, wajib bagi umat beragama dan negara untuk mematuhi hukum Syariah, termasuk memboikot produk-produk pro Israel [7]. Fatwa-fatwa ini mendukung gerakan boikot dan sejalan dengan kampanye BDS (*Boycott, Divestment, and Sanctions*) yang telah digaungkan secara global.

Sosial media seperti X (Twitter) memungkinkan orang-orang untuk secara aktif menyuarakan opini atau aspirasi mereka terkait berbagai isu, salah satunya adalah kampanye boikot produk yang mendukung Israel. Platform ini memungkinkan pengguna untuk menyebarkan informasi, berbagi konten, dan berdiskusi secara *real-time*, sehingga menjadi sarana efektif untuk menggaungkan suara publik [8]. Dalam penelitian Supriadi dan Fatmasari (2021)

menunjukkan bahwa karena jangkauannya yang luas, X berfungsi sebagai representasi opini publik. Dalam konteks ini, Analisis sentimen terhadap data dari X menjadi penting untuk memahami bagaimana opini masyarakat terbentuk [9].

Analisis sentimen adalah teknik yang digunakan untuk menentukan opini atau sentimen yang terkandung dalam suatu teks. Teknik ini digunakan untuk menganalisis berbagai jenis teks, seperti ulasan produk, postingan media sosial, artikel berita, dan lain sebagainya [10]. Teknik ini dapat mengubah data yang tidak terstruktur menjadi informasi yang berharga yang diklasifikasikan ke dalam sentimen positif, negatif, atau netral berdasarkan polaritas emosional yang diberikan [11]. Algoritma *Machine Learning* seperti *Random Forest* dan *Support Vector Machine (SVM)* dapat digunakan untuk mengklasifikasi sentimen. Algoritma SVM bekerja dengan menemukan *hyperplane* yang memaksimalkan margin antara kelas dalam ruang fitur yang berdimensi lebih tinggi [12]. Sedangkan *Random Forest* bekerja dengan menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi dan mengatasi *overfitting* [13].

Penelitian sebelumnya yang dilakukan oleh Alifa dan Alita (2024) mengenai analisis sentimen terhadap boikot produk pro israel yang membandingkan 2 kernel pada algoritma SVM memperoleh akurasi kernel linear sebesar 91% dan rbf sebesar 92% [14]. Penelitian yang dilakukan oleh Fahrani dan Aryanto (2024) mengenai analisis sentimen Konflik Israel dan Palestina menggunakan algoritma SVM dan *Naïve Bayes*, memperoleh kinerja unggul pada SVM dengan akurasi 97% sedangkan akurasi *Naïve Bayes* sebesar 86% [15]. Penelitian yang dilakukan oleh Pratama dkk (2023) mengenai analisis sentimen yang membandingkan kinerja algoritma *Naïve Bayes* dan *Random Forest*, dimana performa *Random Forest* lebih unggul dengan akurasi 85,15% dibandingkan *Naïve Bayes* yang memperoleh akurasi 79,74% [16].

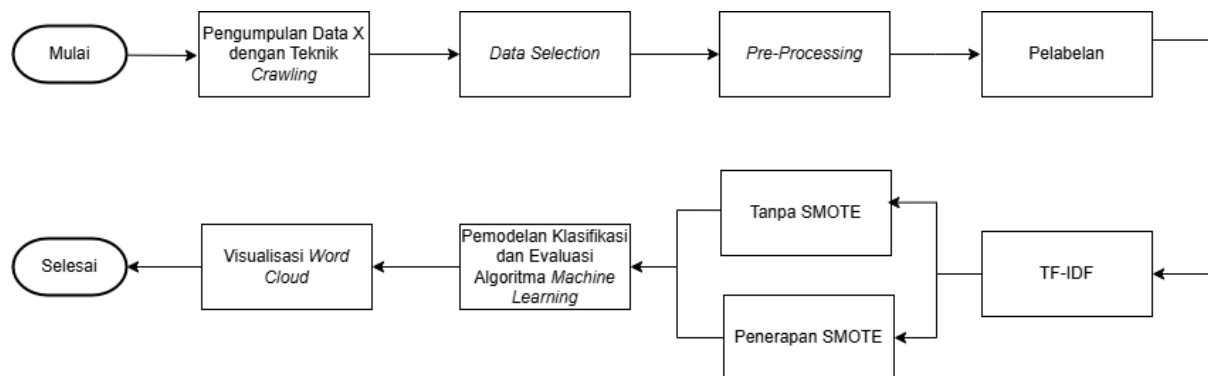
Penelitian yang dilakukan oleh Adriana dkk (2022) mengenai analisis sentimen terhadap aksi demonstrasi, algoritma SVM memiliki kinerja lebih baik dibandingkan *Random Forest* dengan akurasi 80,3%, sedangkan *Random Forest* memiliki akurasi 77,7% [17]. Penelitian yang dilakukan oleh Septiana dan Alita (2024) mengenai analisis sentimen *quick count* pemilu 2024 memperoleh akurasi 78% pada algoritma *Random Forest* dan akurasi 80% pada algoritma SVM [18]. Penelitian yang dilakukan oleh Hartanto dkk (2024) yang membahas penerapan SMOTE pada analisis sentimen terhadap pemindahan Ibu Kota Negara (IKN), memperoleh hasil peningkatan performa algoritma SVM dari 85% menjadi 92% dan algoritma *Random Forest* dari 84% menjadi 91%. Hasil ini menunjukkan bahwa metode SMOTE efektif dalam memperkuat kemampuan algoritma klasifikasi untuk menangani ketidakseimbangan data [19].

Berdasarkan uraian tersebut, pada penelitian ini mengangkat topik Klasifikasi Sentimen Pengguna X Terhadap Pemboikotan Produk Pro Israel. Penelitian juga akan menerapkan SMOTE *Oversampling* untuk mengatasi ketidakseimbangan data. Penelitian ini bertujuan untuk mengklasifikasi sentimen pengguna X terhadap boikot produk pro Israel dan membandingkan performa algoritma SVM dan *Random Forest* dalam klasifikasi sentimen. Hasil penelitian diharapkan dapat memberikan gambaran sentimen masyarakat mengenai opini publik Indonesia terkait isu boikot produk pro Israel.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini mengadopsi pendekatan sistematis yang disajikan mulai dari tahap pengumpulan data hingga tahap visualisasi *word cloud*. Tahapan penelitian dapat diamati pada Gambar 1.



Gambar 1. Metodologi Penelitian

Proses pengumpulan data pada X dilakukan dengan menggunakan teknik *crawling* dengan *tools google collab*, bahasa pemrograman *python*. *Keyword* yang digunakan yaitu “Boikot Produk Pro Israel”, “Boikot Produk Israel”, dan “Boikot Brand Israel” dalam rentang waktu November 2023 – November 2024. Selanjutnya dilakukan proses seleksi data seperti penghapusan data *tweet* yang tidak terkait dengan topik penelitian. Kemudian dilakukan proses *pre-processing* dimulai dari *cleaning* dengan menghapus *noise* pada data, *case folding* mengubah data menjadi *lower case*,

Tokenizing memecah kalimat menjadi kata-kata, *Stopword* penghapusan kata umum yang tidak memberikan makna, serta *stemming* menghapus imbuhan dan mengubah menjadi kata dasar. Setelah dilakukan tahapan *pre-processing*, jumlah data bersih yang digunakan pada penelitian ini adalah 2.275.

Langkah berikutnya dilakukan pelabelan dengan mengkategorikan sentimen menjadi kelas positif, negatif, dan netral oleh secara manual oleh annotator yang memiliki keahlian di bidang Bahasa Indonesia. Kemudian dilakukan pembobotan kata dengan TF-IDF untuk mengetahui seberapa unik dan penting sebuah kata dilakukan dengan menggunakan modul *sklearn* dan bahasa pemrograman *Python*. tahap berikutnya, dilakukan dua pendekatan yaitu tanpa SMOTE dan dengan SMOTE untuk menangani ketidakseimbangan data sebelum dilakukan *training*. Tahapan pemodelan klasifikasi dilakukan dengan algoritma *Machine Learning* yaitu SVM dan *Random Forest*. Tahapan berikutnya evaluasi menggunakan metrik *accuracy*, *precision*, dan *recall*. Langkah terakhir adalah membuat visualisasi *word cloud* dengan menampilkan semua kata dalam dataset yang dikelompokkan berdasarkan kategori sentimen positif, negatif, dan netral.

2.2 Klasifikasi Sentimen

Proses mengidentifikasi secara otomatis orientasi opini dalam sebuah teks dengan mengelompokkannya ke dalam kelas sentimen (positif, negatif, dan netral) dikenal sebagai klasifikasi sentimen [20]. Sentimen adalah pendapat pribadi seseorang tentang suatu hal [21]. Komentar dan sentimen dari masyarakat, baik berupa opini maupun pengalaman pribadi, dapat dikumpulkan dalam jumlah besar dan dianggap sebagai data berharga untuk dianalisis [22]. Dengan memanfaatkan algoritma pembelajaran mesin, klasifikasi sentimen dapat memberikan pemahaman mendalam tentang persepsi publik terhadap suatu isu, produk, atau peristiwa serta mengidentifikasi tren opini publik secara *real-time* [23].

2.3 Support Vector Machine

Algoritma *Support Vector Machine* (SVM) adalah salah satu teknik pembelajaran mesin yang efisien untuk mengelompokkan data berdasarkan teori pembelajaran statistik. SVM membagi data ke dalam kelas-kelas yang berbeda dengan mengidentifikasi *hyperplane* terbaik dalam ruang fitur berdimensi tinggi, di mana model klasifikasi yang unggul dihasilkan oleh *hyperplane* yang memaksimalkan margin antara kelas data [24]. Dalam konteks sentimen analisis, kelas-kelas ini bisa berupa sentimen positif, negatif, atau netral. Kelebihan dari algoritma SVM yaitu memiliki kemampuan untuk mengubah data nonlinear menjadi linear dengan menggunakan fungsi kernel, sehingga dapat memisahkan data yang kompleks dengan efisiensi tinggi [25]. Dalam algoritma SVM, pemilihan kernel sangat penting karena menentukan ruang fitur di mana fungsi pengklasifikasi akan dibangun. Teknik kernel trick memungkinkan pemilihan jenis kernel yang tepat, seperti *linear*, *RBF*, *polynomial*, atau *sigmoid*, agar dapat disesuaikan dengan berbagai jenis data [26]. Berikut ini rumus dari persamaan kernel SVM [27]:

Kernel *linear* :

$$K(x, y) = x \cdot y \tag{1}$$

Kernel *rbf*

$$K(x, y) = e^{-|x \cdot y|^2 / 2\sigma^2} \tag{2}$$

Kernel *sigmoid*

$$(x, y) = \tanh(Kx \cdot y - \delta) \tag{3}$$

Kernel *polynomial*

$$K(x, y) = (x \cdot y + 1)^p \tag{4}$$

Untuk meningkatkan kinerja model dapat dilakukan optimasi parameter. Adapun penggunaan parameter dan nilainya disajikan pada Tabel 1.

Tabel 1. Optimasi Parameter SVM

Parameter	Nilai	Penerapan Kernel
Cost (C)	0.001, 0.01, 0.1, 1, 10, 100	<i>Linear</i> , <i>rbf</i> , <i>sigmoid</i> , dan <i>polynomial</i>
Gamma (γ)	0.001, 0.01, 0.1, 1, 10, 100	<i>Rbf</i> , <i>sigmoid</i> , dan <i>polynomial</i>
Coefficient (coef)	0, 1, 5, 8, 10	<i>Sigmoid</i> dan <i>polynomial</i>
Degree (d)	2, 3, 4, 5	<i>Polynomial</i>

Pemberian rentang nilai pada parameter tersebut didasarkan pada penelitian yang dilakukan Tuttle dkk (2020) [28]. Pemilihan nilai parameter dapat memberikan pengaruh yang besar terhadap hasil klasifikasi model.

2.4 Random Forest

Random Forest merupakan *ensemble learning* yang menggabungkan banyak pohon keputusan. Setiap pohon dilatih pada sampel data yang berbeda secara acak, menggunakan subset fitur yang dipilih secara acak pula. *Random Forest* mengambil keputusan berdasarkan suara mayoritas dari semua pohon keputusan, kelas yang paling banyak diprediksi

oleh pohon-pohon akan menentukan hasil akhir [13]. Algoritma *Random Forest* dipilih karena kemampuan dalam melakukan klasifikasi dan regresi dengan akurasi yang tinggi. *Random Forest* unggul dalam menangani data yang tidak lengkap, *outlier*, dan memiliki mekanisme seleksi fitur yang efektif untuk meningkatkan akurasi. *Random Forest* juga cukup mudah digunakan dan hanya memerlukan beberapa pengaturan parameter [29]. Adapun rumus persamaan algoritma *Random Forest* adalah sebagai berikut:

$$y = \text{mode} \{h_1(x), h_2(x), \dots, h_n(x)\} \tag{5}$$

Berdasarkan persamaan tersebut, y adalah prediksi akhir (kelas mayoritas), $h_i(x)$ adalah prediksi *decision tree* ke- i untuk input x dan n adalah jumlah total *decision tree* dalam *Random Forest*. Untuk meningkatkan performa model pada *Random Forest* dapat dilakukan optimasi parameter. Berikut ini penggunaan parameter pada *Random Forest* dapat dilihat pada Tabel 2.

Tabel 2. Optimasi Prameter *Random Forest*

Parameter	Nilai
<i>n_estimators</i>	50, 100, 200, 400
<i>max_depth</i>	10, 20, 40, None
<i>min_samples_split</i>	2, 5, 10, 20
<i>min_samples_leaf</i>	1, 2, 4, 8
<i>bootstrap</i>	True, false

Pemberian nilai parameter tersebut didasarkan pada penelitian yang dilakukan oleh Suryadi dkk (2024) [30]. Meskipun *Random Forest* biasanya bekerja dengan baik dengan pengaturan *default*, pengoptimalan parameter dapat membantu meningkatkan akurasi dan mengurangi *overfitting*.

2.5 SMOTE

Masalah ketidakseimbangan kelas pada dataset dapat diatasi dengan teknik *oversampling*. Salah satu penerapan metode *oversampling* yaitu SMOTE (*Synthetic Minority Over-sampling Technique*). SMOTE merupakan teknik yang populer untuk mengatasi masalah ketidakseimbangan kelas, yang dapat mempengaruhi kinerja algoritma *machine learning*. Untuk membantu menyeimbangkan dataset, metode ini menghasilkan data sintetis untuk kelas minoritas. Sebagai contoh, Jika memiliki dataset dengan 100 contoh, 10 di antaranya adalah minoritas, SMOTE akan menghasilkan contoh tambahan untuk kelas minoritas agar sama dengan kelas mayoritas [31].

3. HASIL DAN PEMBAHASAN

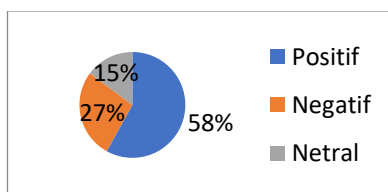
Penelitian ini menggunakan *tweet* dan komentar dari pengguna X dalam rentang waktu 1 tahun terakhir (November 2023 - November 2024). Data tersebut akan diolah dan dianalisis guna memahami sentimen masyarakat terhadap kampanye boikot produk Israel pada platform tersebut. Kemudian dilakukan *data selection* dengan menghapus *tweet* atau komentar yang tidak relevan dengan topik penelitian. Maka diperoleh data selama rentang waktu tersebut berjumlah 2275. Data tersebut diproses dengan *pre-processing* yang terdiri dari tahapan *cleaning*, *case folding*, *tokenizing*, *stopword*, *stemming*. Hasil *pre-processing* dapat dilihat pada Tabel 3.

Tabel 3. Pre-Processing

Proses	Hasil
Data Mentah	Boikot terus sampai benar benar stop kasih sponsor ke Israel. #BoikotprodukIsrael
<i>Cleaning</i>	Boikot terus sampai benar benar stop kasih sponsor ke Israel
<i>Case Folding</i>	boikot terus sampai benar benar stop kasih sponsor ke israel
<i>Tokenizing</i>	['boikot', 'terus', 'sampai', 'benar', 'benar', 'stop', 'kasih', 'sponsor', 'ke', 'israel']
<i>Stopword</i>	['boikot', 'stop', 'kasih', 'sponsor', 'israel']
<i>Stemming</i>	boikot stop kasih sponsor israel

3.1 Pelabelan

Pelabelan bertujuan untuk memberikan nilai positif, negatif, atau netral pada setiap komentar. Proses pelabelan ini dilakukan secara manual oleh annotator Bahasa Indonesia. Adapun hasil pelabelan dapat dilihat pada Gambar 2.



Gambar 2. Pelabelan

Berdasarkan hasil pelabelan yang dilakukan, pada platform X mayoritas komentar yang ditemukan bernada positif yaitu 1321 (58%) komentar. Hal ini menunjukkan dukungan yang kuat terhadap gerakan boikot ini. Adapun jumlah komentar negatif sebesar 617 (27%) komentar, mencerminkan adanya perbedaan pendapat atau perdebatan mengenai isu ini. Sementara itu, terdapat pula sejumlah komentar netral yaitu 337 (15%) komentar yang tidak secara eksplisit menyatakan dukungan atau penolakan.

3.2 TFIDF

Tahapan selanjutnya adalah penerapan teknik TF-IDF untuk memberikan pembobotan pada setiap istilah dalam dokumen. Nilai TF-IDF yang tinggi mengindikasikan bahwa suatu kata sangat relevan dengan dokumen tersebut dan dapat digunakan untuk membedakan dokumen tersebut dengan dokumen lainnya. Adapun hasil pada proses tahapan ini dapat dilihat pada Tabel 4.

Tabel 4. Hasil TFIDF

No	boikot	dampak	israel	palestina	produk	...	zionis
1	0,08	0.0	0.09	0.2	0.0		0.0
2	0.06	0.0	0.06	0.0	0.09		0.0
3	0.12	0.0	0.06	0.0	0.9		0.0
....
2.275	0.11	0.16	0.07	0.0	0.0		0.0

3.3 Penerapan SMOTE

Dalam penelitian ini, data untuk melatih model menunjukkan ketidakseimbangan antar kelas. Untuk mengatasi hal ini, diterapkan metode SMOTE untuk menyeimbangkan jumlah data pada kelas minor. Hasil penerapan SMOTE dapat dilihat pada Tabel 5

Tabel 5. Penerapan SMOTE

Penerapan	Positif	Negatif	Netral	Total Data
Tanpa SMOTE	1321	617	337	2275
SMOTE	1321	1321	1321	3963

Dari tabel tersebut dapat dilihat bahwa proporsi kategori negatif dan netral secara signifikan lebih sedikit daripada kategori positif. Setelah menggunakan SMOTE, jumlah data menjadi sama yaitu 1.321 data untuk masing-masing kategori positif, negatif, dan netral. Adapun total data setelah penerapan SMOTE sebesar 3.963. Penerapan SMOTE membantu meningkatkan representasi setiap kelas dalam dataset dengan menghasilkan data sintesis untuk kelas-kelas minoritas.

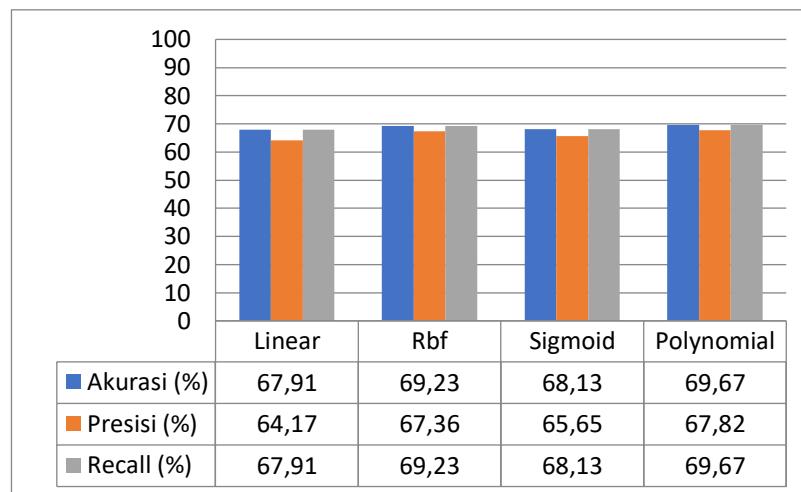
3.4 Hasil Performa SVM

Proses pembagian data dilakukan menggunakan *hold-out validation* dengan rasio data *training* dan data *testing* 80:20. Kemudian dilakukan klasifikasi dengan algoritma SVM dan evaluasi menggunakan metrik seperti *Accuracy*, *precision*, dan *recall*. Pada penelitian ini dilakukan pengujian masing-masing kernel yaitu kernel *linear*, *rbf*, *sigmoid*, dan *polynomial*. Untuk mengoptimalkan pengaturan parameter maka diterapkan konsep *Grid Search*. Beberapa kombinasi parameter seperti nilai *cost*, *gamma*, *coefficient* dan *degree* digunakan dalam proses ini untuk menemukan pengaturan parameter yang memberikan kinerja terbaik. Nilai parameter yang digunakan dalam pengujian ini didasarkan pada optimasi parameter SVM yang tercantum dalam Tabel 1. Adapun hasil performa klasifikasi algoritma SVM disajikan pada Tabel 6.

Tabel 6. Hasil Performa SVM

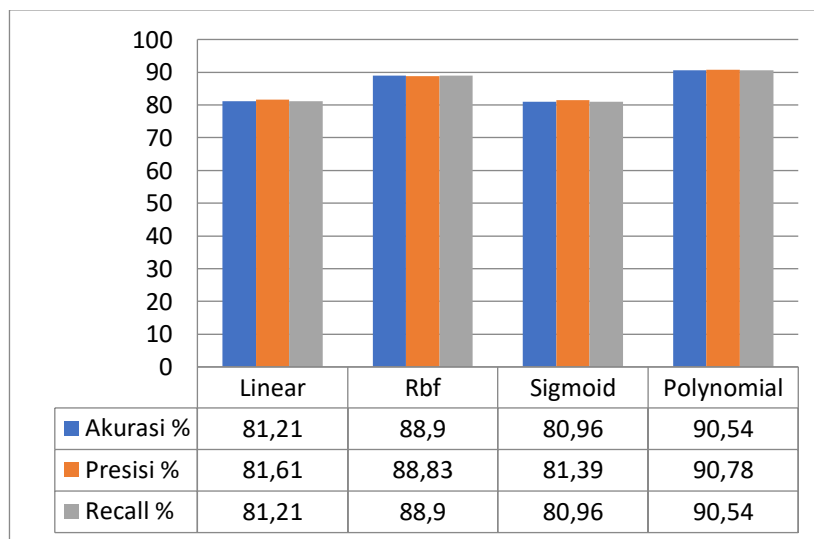
Kategori	Kernel	C	γ	coef	d	Accuracy	Precision	Recall
Tanpa SMOTE	Linear	1	-	-	-	67,91%	64,17%	67,91%
	Rbf	100	1	-	-	69,23%	67,36%	69,23%
	Sigmoid	1	1	0	-	68,13%	65,65%	68,13%
	Polynomial	0,001	10	5	3	69,67%	67,82%	69,67%
SMOTE	Linear	10	-	-	-	81,21%	81,61%	81,21%
	Rbf	100	1	-	-	88,9%	88,83%	88,9%
	Sigmoid	100	0,1	0	-	80,96%	81,39%	80,96%
	Polynomial	0,001	10	0	3	90,54%	90,78%	90,54%

Berdasarkan tabel tersebut, terlihat hasil yang menunjukkan bahwa kinerja SVM sangat dipengaruhi oleh interaksi antara pemilihan kernel, nilai parameter dan penerapan teknik SMOTE. Kombinasi yang tepat dapat meningkatkan akurasi klasifikasi secara signifikan. Adapun tampilan visualisasi dari performa SVM tanpa SMOTE dan dengan penerapan SMOTE pada masing-masing kernel dapat dilihat pada Gambar 3 dan Gambar 4.



Gambar 3. Performa SVM tanpa SMOTE

Grafik ini menunjukkan bahwa kernel *polynomial* dan *rbf* memberikan hasil klasifikasi yang paling baik, terutama dalam hal akurasi sebesar 69,67% dan 69,23%. Kernel *sigmoid* juga cukup baik dengan akurasi 68,13%, namun kernel *linear* memiliki kinerja terendah dengan akurasi 67,91%.



Gambar 4. Performa SVM menggunakan SMOTE

Grafik di atas menunjukkan perbandingan akurasi, presisi, dan recall untuk berbagai jenis kernel yang digunakan dalam algoritma SVM dengan menerapkan konsep SMOTE. Kernel *polynomial* memberikan performa terbaik untuk ketiga metrik dengan akurasi sebesar 90,54%, diikuti akurasi kernel *rbf* sebesar 88,9%. Kernel *linear* dan kernel *sigmoid* memiliki performa yang kurang, dimana akurasi kernel linear 81,21% dan akurasi kernel sigmoid adalah 80,96%.

Kedua grafik tersebut menunjukkan bahwa penggunaan SMOTE secara signifikan meningkatkan nilai akurasi, presisi, dan recall untuk setiap kernel SVM. Tanpa penerapan SMOTE, kinerja model cenderung lebih rendah akibat ketidakseimbangan data. Kernel *Polynomial* memiliki kinerja terbaik dalam kedua scenario berkat kemampuannya dalam menangkap pola non-linear yang kompleks. Kernel *Polynomial* memungkinkan SVM untuk mengatasi masalah data yang tidak linier dengan cara memetakan data ke ruang yang lebih cocok, sehingga meningkatkan akurasi.

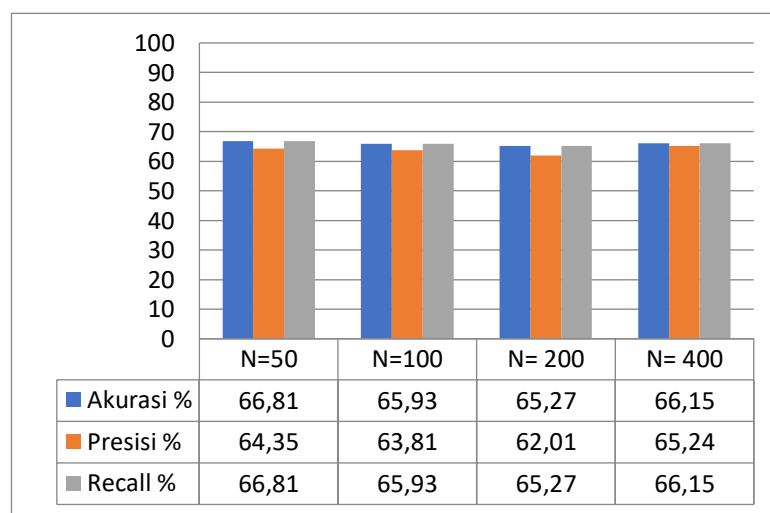
3.5 Hasil Performa *Random Forest*

Pada tahap ini dilakukan klasifikasi menggunakan algoritma *Random Forest*. Untuk memastikan model ini bekerja secara optimal, diperlukan pengaturan parameter yang tepat. Konsep *Grid Search* juga diterapkan pada model ini agar dapat mengoptimalkan parameter dengan nilai terbaik. Kombinasi parameter yang digunakan pada algoritma ini yaitu *n_est* (banyaknya jumlah pohon), *max_d* (kedalaman maksimum dari pohon), *min_sp* (jumlah sampel yang diperlukan untuk memisahkan node), *min_sl* (jumlah sampel yang ada di daun akhir pohon), dan *bootstrap* (apakah sampel *bootstrap* digunakan dalam membangun pohon). Nilai parameter yang digunakan dalam pengujian ini didasarkan pada optimasi parameter *Random Forest* yang tercantum dalam Tabel 2. Pada penelitian ini lakukan uji jumlah pohon dengan nilai 50, 100, 200, dan 400. Hasil klasifikasi menggunakan *Random Forest* dapat dilihat pada Tabel 7.

Tabel 7. Hasil Performa Random Forest

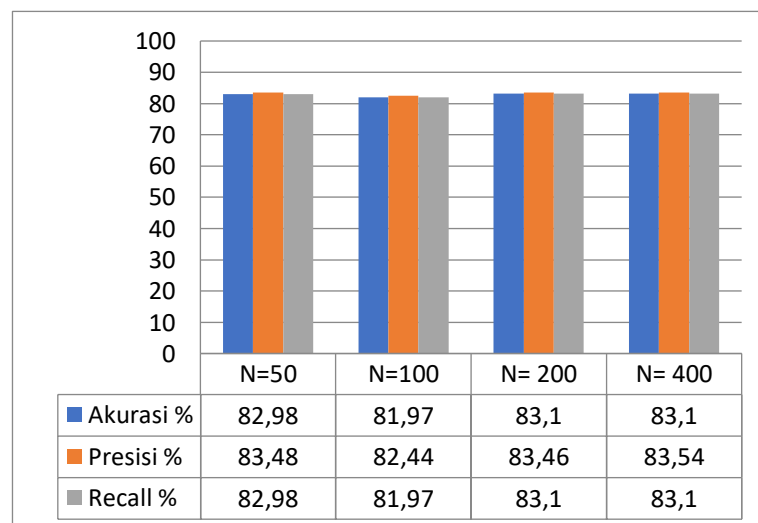
Kategori	<i>N_{est}</i>	<i>Max_d</i>	<i>Min_{sp}</i>	<i>Min_{sl}</i>	<i>Boot</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
Tanpa SMOTE	50	None	10	1	False	66,81%	64,35%	66,81%
	100	None	5	1	True	65,93%	63,81%	65,93%
	200	None	2	1	True	65,27%	62,01%	65,27%
	400	40	10	2	False	66,15%	65,24%	66,15%
SMOTE	50	None	2	1	False	82,98%	83,48%	82,98%
	100	None	5	1	False	81,97%	82,44%	81,97%
	200	None	2	1	False	83,1%	83,46%	83,1%
	400	None	2	1	False	83,1%	83,54%	83,1%

Berdasarkan Tabel 7, terlihat hasil yang menunjukkan bahwa pengaturan parameter dan penerapan teknik SMOTE mempengaruhi performa algoritma Random Forest. Adapun tampilan visualisasi dari performa *Random Forest* tanpa SMOTE dan dengan penerapan SMOTE pada masing-masing jumlah pohon yang diuji dapat dilihat pada gambar 5 dan gambar 6.



Gambar 5. Performa *Random Forest* tanpa SMOTE

Grafik tersebut menunjukkan bahwa pada N=50 menghasilkan nilai akurasi dan *recall* 66.81%, nilai presisinya 64,35%. Pada N=100 nilai akurasi dan *recall* sebesar 65,93%, nilai presisi 63,81%. Sementara N=200 menghasilkan sedikit penurunan nilai akurasi dan *recall* yaitu 65.27%, dan nilai presisi 62,01%. Pada jumlah pohon dinaikkan menjadi N=400 menghasilkan peningkatan nilai akurasi dan *recall* yaitu 66.15%, dan nilai presisinya 65,24. Pada model ini nilai akurasi tertinggi diperoleh pada jumlah pohon N=50.



Gambar 6. Performa *Random Forest* menggunakan SMOTE

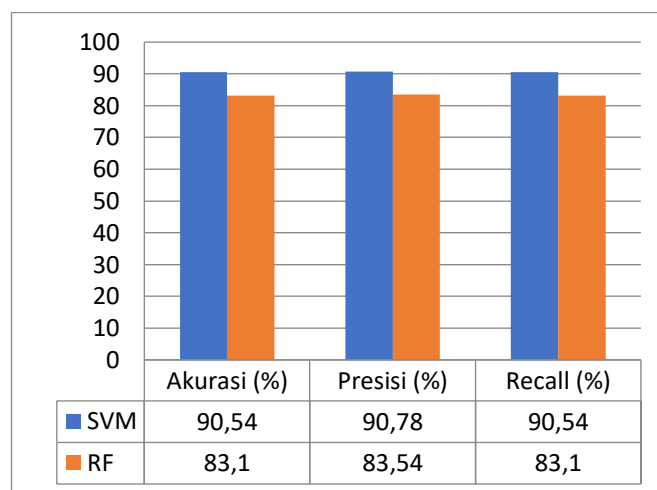
Berdasarkan grafik pada Gambar 6, pada N=50 akurasinya sebesar 82,98% dan presisinya 83,48%. Ketika jumlah sampel ditingkatkan menjadi N=100, model menunjukkan sedikit penurunan akurasi dan *recall* sebesar

81,97% dan presisi sebesar 82,44%. Namun, ketika jumlah sampel ditingkatkan menjadi N=200 dan N=400, model menunjukkan kinerja yang stabil dengan akurasi dan *recall* sebesar 83,1% dan sedikit peningkatan presisi pada N=400 menjadi 83,54%.

Dari kedua grafik tersebut, algoritma Random Forest tanpa SMOTE memiliki akurasi, presisi, dan recall berada pada kisaran 62%-66%. Sebaliknya, penerapan SMOTE meningkatkan ketiga metrik secara signifikan hingga mencapai kisaran 81%-83%, menandakan keberhasilan teknik ini dalam memperbaiki distribusi data dan membuat model mampu mengenali pola dari kelas minoritas dengan lebih baik.

3.6 Perbandingan Performa SVM dan Random Forest

Berdasarkan klasifikasi yang telah dilakukan, akurasi dengan nilai terbesar pada algoritma *support vector machine* diperoleh saat menerapkan teknik SMOTE yaitu 90,54% dengan kernel *polynomial*. Pada algoritma *Random Forest* akurasi terbesar juga diperoleh saat menggunakan teknik SMOTE yaitu 83,1% untuk jumlah pohon 200 dan 400 dengan nilai presisi yang juga meningkat pada jumlah pohon 400. Pada penelitian ini teknik SMOTE efektif digunakan untuk meningkatkan performa model dan mengatasi ketidakseimbangan data. Gambar perbandingan kedua algoritma dapat dilihat pada Gambar 7.



Gambar 7. Perbandingan algoritma SVM dan *Random Forest*

Dari perbandingan tersebut, Algoritma SVM memiliki performa yang lebih unggul dibandingkan *Random Forest*. Kinerja SVM unggul dapat dikaitkan dengan kemampuannya memaksimalkan margin pemisahan kelas serta menangani pola *non-linear* melalui kernel, yang efektif untuk dataset ini.

3.7 Visualisasi Word Cloud

Tahapan berikutnya adalah melakukan visualisasi menggunakan *word cloud*. Proses ini menampilkan kata yang muncul pada setiap sentimen. Pada sentimen positif, *word cloud* ini menyoroti dukungan terhadap ajakan boikot produk terkait dengan Israel. Kemunculan Kata dengan 5 frekuensi tertinggi seperti “boikot”, “lanjut,” “dukung”, “Palestina,” dan “aksi” dalam *word cloud* sentimen positif mencerminkan dukungan kuat terhadap gerakan pemboikotan produk pro-Israel, dengan fokus pada solidaritas, keberlanjutan aksi, dan keterlibatan nyata dalam mendukung perjuangan Palestina. Hasil visualisasi *word cloud* sentimen positif dapat dilihat pada Gambar 8.



Gambar 8. *Word Cloud* Positif

Pada sentimen negatif, *Word cloud* ini menyoroti penolakan terhadap gerakan boikot. Kemunculan Kata dengan 5 frekuensi tertinggi seperti “salah”, “kena”, “kerja”, “usaha”, dan “Pakai” mencerminkan kekhawatiran



- 83-Tahun-2023-tentang-Hukum-Dukungan-Terhadap-Perjuangan-Palestina.pdf
- [7] Amin Awal Amarudin, Novi Ria Ananta, Nurun Nisaul Khusna, Regita Juninda Berliani, and Sri Oktavianah, “Analisis Literasi Halal Dan Preferensi Produk Yang Diboikot Pada Mahasiswa Universitas KH. A. Wahab Hasbullah,” *Pop. J. Penelit. Mhs.*, vol. 3, no. 1, pp. 210–222, 2024, doi: 10.58192/populer.v3i1.1948.
 - [8] D. Kundana and Chairani, “Data-Driven Analysis of Borobudur Ticket Sentiment Using Naïve Bayes,” *APTISI Trans. Technopreneursh.*, vol. 5, no. 2Sp, pp. 221–233, 2023, doi: 10.34306/att.v5i2sp.353.
 - [9] Apif Supriadi and Fatmasari, “Implementasi Metode Klasifikasi Naive Bayes Pada Sistem Analisis Opini Pengguna Twitter Berbasis Web,” *J. Sist. Inf.*, vol. 10, no. 1, pp. 46–54, 2021, doi: 10.51998/jsi.v10i1.356.
 - [10] J. K. Kim, A. Khondker, M. E. Chua, M. Rickard, and A. Lorenzo, “Sentiment analysis of U.S. News & World Report Best Children’s Hospital urology rankings: A difference in positivity between the public and academic worlds,” *J. Pediatr. Urol.*, vol. 20, pp. S81–S85, 2024, doi: 10.1016/j.jpuro.2024.06.001.
 - [11] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, “Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions,” *Heliyon*, vol. 10, no. 9, p. e27863, 2024, doi: 10.1016/j.heliyon.2024.e27863.
 - [12] R. Zhang, Y. Li, Y. Gui, D. J. Armaghani, and M. Yari, “A stacked multiple kernel support vector machine for blast induced flyrock prediction,” *Geohazard Mech.*, vol. 2, no. 1, pp. 37–48, 2024, doi: 10.1016/j.ghm.2024.01.002.
 - [13] R. Sheila, T. Rahmayani, and F. Budiman, “Analisa Optimasi Grid Search pada Algoritma Random Forest dan Decision Tree untuk Klasifikasi Stunting,” vol. 6, no. 3, pp. 1537–1546, 2024, doi: 10.47065/bits.v6i3.6128.
 - [14] C. F. Alifa and D. Alita, “Analisis Opini Publik Tentang Boikot Produk Pro-Israel di Twitter Berbahasa Indonesia Menggunakan Metode SVM,” *J. Inform. J. Pengemb. IT*, vol. 9, no. 2, pp. 112–120, 2024, doi: 10.30591/jpit.v9i2.6559.
 - [15] F. Fahrani, J. Aryanto, and U. T. Yogyakarta, “Sentiment Analysis of Public Opinion on the Palestinian-Israeli Conflict using Support Vector Machine and Naive Bayes Algorithms,” *JSRET (Journal Sci. Res. Educ. Technol.*, vol. 3, no. 4, pp. 1890–1900, 2024, doi: <https://doi.org/10.58526/jsret.v3i4.606>.
 - [16] M. L. Pratama, Y. V. Via, and E. P. Mandyartha, “Analisis Performansi Naive Bayes Dan Random Forest Terhadap Sentimen Kenaikan Harga BBM di Indonesia,” *Scan J. Teknol. Inf. dan Komun.*, vol. 18, no. 1, 2023, doi: 10.33005/scan.v18i1.3837.
 - [17] Ni Made Tara Okta Adriana, I Made Agus Dwi Suarjaya, and Dwi Putra Githa, “Analisis Sentimen Publik Terhadap Aksi Demonstrasi di Indonesia Menggunakan Support Vector Machine Dan Random Forest,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, pp. 257–267, 2023, doi: 10.51454/decode.v3i2.187.
 - [18] I. Septiana and D. Alita, “Perbandingan Random Forest dan SVM dalam Analisis Sentimen Quick Count Pemilu 2024,” vol. 9, no. 3, pp. 224–233, 2024, doi: 10.30591/jpit.v9i3.6640.
 - [19] A. Almas, U. D. Zd, and A. D. W. I. Hartanto, “Analisis Sentimen Pemindahan Ibu Kota Negara (IKN) Menggunakan Metode Oversampling Synthetic Minority (SMOTE),” vol. 9, pp. 324–335, 2024, doi: <https://doi.org/10.24252/instek.v9i2.50944>.
 - [20] A. A. Syed, F. L. Gaol, A. Boediman, and W. Budiharto, “Airline reviews processing: Abstractive summarization and rating-based sentiment classification using deep transfer learning,” *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, p. 100238, 2024, doi: 10.1016/j.jjime.2024.100238.
 - [21] E. Suryati, Styawati, and A. A. Aldino, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM),” *J. Teknol. Dan Sist. Inf.*, vol. 4, no. 1, pp. 96–106, 2023, doi: <https://doi.org/10.33365/jtsi.v4i1.2445>.
 - [22] Ash Shiddicky and Surya Agustian, “Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 pada Media Sosial Twitter menggunakan Metode Logistic Regression,” *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 3, no. 2, pp. 99–106, 2022, doi: 10.37859/coscitech.v3i2.3836.
 - [23] A. Daza, N. D. González Rueda, M. S. Aguilar Sánchez, W. F. Robles Espiritu, and M. E. Chauca Quiñones, “Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis and Systematic Literature Review, Challenges and Future Works,” *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, 2024, doi: 10.1016/j.jjime.2024.100267.
 - [24] J. S. Pimentel, R. Ospina, and A. Ara, “A novel fusion Support Vector Machine integrating weak and sphere models for classification challenges with massive data,” *Decis. Anal. J.*, vol. 11, no. March, p. 100457, 2024, doi: 10.1016/j.dajour.2024.100457.
 - [25] T. B. Rizal Amegia Saputra, Diah Puspitasari, “Deteksi Kematangan Buah Melon dengan Algoritma Support Vector Machine Berbasis Ekstraksi Fitur GLCM,” *J. Infortech*, vol. 4, no. 2, pp. 200–206, 2022.
 - [26] D. L. Rianti, Y. Umaidah, and A. Voutama, “Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel Support Vector Machine,” *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 6, no. 1, p. 98, 2021, doi: 10.30998/string.v6i1.9993.
 - [27] C. Sonjaya, A. F. Nur Masruriyah, D. Sulisty Kusumaningrum, and A. Rizky Pratama, “The Performance Comparison of Classification Algorithm in Order to Detecting Heart Disease,” *Intern. (Information Syst. Journal)*, vol. 5, no. 2, pp. 166–175, 2022, doi: 10.32627/internal.v5i2.595.
 - [28] J. F. Tuttle, L. D. Blackburn, and K. M. Powell, “On-line classification of coal combustion quality using nonlinear SVM for improved neural network NOx emission rate prediction,” *Comput. Chem. Eng.*, vol. 141, p. 106990, 2020, doi: 10.1016/j.compchemeng.2020.106990.
 - [29] N. A. Hapsari and A. D. Indriyanti, “Analisis Sentimen pada Aplikasi Dompot Digital Menggunakan Algoritma Random Forest,” *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 04, no. 03, pp. 186–192, 2023, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/JEISBI/article/view/55696>
 - [30] M. K. Suryadi, R. Herteno, S. W. Saputro, M. R. Faisal, and R. A. Nugroho, “A Comparative Study of Various Hyperparameter Tuning on Random Forest Classification with SMOTE and Feature Selection Using Genetic Algorithm in Software Defect Prediction,” *J. Electron. Electromed. Eng. Med. Informatics*, vol. 6, no. 2, pp. 137–147, 2024, doi: 10.35882/jeeemi.v6i2.375.
 - [31] P. S. Yadav, R. S. Rao, A. Mishra, and M. Gupta, “Ensemble methods with feature selection and data balancing for improved code smells classification performance,” *Eng. Appl. Artif. Intell.*, vol. 139, no. PA, p. 109527, 2025, doi: 10.1016/j.engappai.2024.109527.