

Deteksi Dark Patterns Biaya Layanan E-Commerce Berdasarkan Perspektif Konsumen Menggunakan Algoritma Support Vector Machine

Divana Taricha Salmalina¹, Khothibul Umam^{1*}, Maya Rini Handayani²

¹Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo, Semarang, Indonesia

²Komunikasi Penyiaran Islam, Fakultas Dakwah dan Komunikasi, UIN Walisongo, Semarang, Indonesia

Email: ¹2208096034@student.walisongo.ac.id, ^{2*}khothibul_umam@walisongo.ac.id, ³maya@walisongo.ac.id

Email Penulis Korespondensi: khothibul_umam@walisongo.ac.id*

Submitted: 27/04/2025; Accepted: 19/05/2025; Published: 30/06/2025

Abstrak– Perkembangan industri e-commerce di Indonesia belakangan ini dihadapkan pada fenomena meningkatnya keluhan konsumen terkait kebijakan biaya layanan yang dinilai kurang transparan, termasuk indikasi adanya praktik dark patterns. Penelitian ini bertujuan mengkaji persepsi konsumen terhadap isu tersebut melalui pendekatan analisis sentimen berbasis machine learning dan deteksi pola manipulatif. Data penelitian diperoleh dari ulasan pengguna di platform media sosial X yang kemudian diproses melalui serangkaian tahapan text mining meliputi pembersihan data, tokenisasi, stopword removal, dan stemming. Analisis sentimen menggunakan algoritma Support Vector Machine (SVM) menunjukkan hasil yang signifikan, dimana 55-78% ulasan di ketiga platform e-commerce (Shopee, Tokopedia, Lazada) tergolong negatif. Analisis TF-IDF mengidentifikasi kata kunci seperti "biaya", "layan" (layanan), dan "mahal" sebagai istilah paling dominan dalam ulasan negatif. Model SVM menunjukkan performa yang cukup baik dengan akurasi mencapai 87% dalam mengklasifikasikan sentimen negatif. Lebih lanjut, analisis tematik terhadap ulasan negatif berhasil mengidentifikasi indikasi dark patterns, khususnya dalam kategori hidden costs (biaya tersembunyi) dan sneak into basket (penambahan produk tanpa disadari) yang muncul secara konsisten di semua platform. Temuan ini tidak hanya mengkonfirmasi adanya pola manipulatif yang berulang dalam industri e-commerce Indonesia, tetapi juga menegaskan urgensi bagi para pelaku industri untuk meningkatkan transparansi dalam kebijakan biaya. Secara praktis, hasil penelitian ini dapat menjadi bahan pertimbangan penting bagi regulator dalam merumuskan kebijakan perlindungan konsumen di era digital yang lebih komprehensif.

Kata Kunci: Biaya Layanan; Dark Pattern; E-commerce; SVM

Abstract– The rapid growth of e-commerce in Indonesia has been accompanied by increasing consumer complaints regarding opaque service fee policies and potential dark pattern practices. This study examines consumer perspectives through sentiment analysis and dark pattern detection. User reviews from social media platform X were processed using text mining techniques (data cleaning, tokenization, stemming) and analyzed with a Support Vector Machine (SVM) algorithm. Results revealed an overwhelming dominance of negative sentiment (55-78% across platforms), with TF-IDF weighting highlighting "fee," "service," and "expensive" as key terms. The SVM model achieved high accuracy (up to 87%) in detecting negative sentiment. Thematic analysis of negative reviews consistently identified two dark pattern categories: hidden costs and sneak-into-basket tactics, suggesting systemic manipulative practices. These findings underscore critical transparency gaps in Indonesia's e-commerce sector and emphasize the need for industry self-regulation to eliminate deceptive designs, and stronger regulatory frameworks for digital consumer protection. The study contributes methodologically by demonstrating SVM's effectiveness in large-scale sentiment analysis of Indonesian e-commerce discourse, while practically informing policy discussions about ethical platform design. Future research should explore cultural factors influencing consumer perception of dark patterns in Southeast Asian markets.

Keywords: Service Fee; Dark patterns; E-commerce; SVM

1. PENDAHULUAN

Perkembangan teknologi dan internet yang semakin pesat di Indonesia memberikan dampak yang besar terhadap perkembangan dunia usaha[1]. Semenjak pandemi Covid-19 yang melanda, perkembangan teknologi digital telah membawa transformasi besar dalam pola konsumsi masyarakat, terutama dalam aktivitas transaksi melalui platform e-commerce. Platform ini merupakan bentuk aktivitas bisnis berbasis digital yang mencakup produk dan layanan secara menyeluruh. Tidak hanya sebatas transaksi jual beli secara daring, e-commerce juga mencakup seluruh rangkaian proses bisnis secara online, mulai dari tahap perencanaan, promosi, penjualan, distribusi, layanan pelanggan, hingga sistem pembayaran[2]. Perkembangan e-commerce di Indonesia cukup pesat, seperti Tokopedia yang muncul sebagai salah satu pelopor e-commerce di Indonesia yang berdiri pada tahun 2009. Kemudian pada tahun 2012 Lazada masuk ke Indonesia dan disusul oleh Shopee pada tahun 2015[3]. Dalam operasional e-commerce, undang-undang yang mencakup aspek perlindungan konsumen, privasi data, dan ketentuan pajak menjadi keharusan untuk dipatuhi pelaku usaha sehingga terhindar dari sanksi hukum, denda finansial, dan kerusakan reputasi[4]. Seiring dengan persaingan yang semakin ketat antar platform e-commerce untuk memberikan fasilitas dan layanan terbaik guna menarik pengguna dan meraih keuntungan yang tinggi, muncul berbagai strategi dan kebijakan yang diterapkan demi mencapai tujuan tersebut. Namun, terdapat indikasi tidak semua kebijakan berjalan transparan dan adil bagi konsumen. Muncul persoalan serius terkait transparansi

informasi mengenai kebijakan biaya layanan yang dibebankan kepada pengguna. Adanya biaya layanan tambahan, seperti biaya transaksi atau biaya administrasi, yang jumlah dan dasar perhitungannya tidak jelas kerap didapati oleh pengguna[5].

Dalam beberapa kasus, biaya layanan yang tidak transparan ini baru terungkap pada tahap akhir transaksi atau dikenal dengan hidden cost. Selain itu, biaya atau item juga muncul secara tidak terduga dengan label “bonus” atau “wajib” tanpa persetujuan pengguna yang dikenal dengan sneak into basket[6]. Fenomena ini merupakan jenis dari dark patterns, yaitu strategi desain antarmuka digital yang bersifat manipulatif dan mengarahkan pengguna untuk mengambil keputusan yang tidak sepenuhnya rasional maupun terinformasi[7]. Dark patterns sering diterapkan oleh platform e-commerce untuk memengaruhi perilaku dan proses pengambilan keputusan pengguna, dengan sering mengabaikan standar etika yang berlaku[8].

Ketidakterbukaan ini tidak hanya menghambat kemampuan konsumen dalam membuat keputusan, tetapi juga berpotensi menyesatkan persepsi terhadap harga awal yang pada akhirnya merugikan konsumen. Dampaknya tidak hanya merugikan konsumen secara finansial, tetapi juga menimbulkan dampak psikologis berupa rasa tertipu, frustrasi, dan hilangnya kepercayaan terhadap platform e-commerce. Ketidakpuasan pengguna kemudian dicurahkan pada ulasan-ulasan di media sosial seperti platform X (Twitter). Ulasan-ulasan ini menjadi sumber data penting untuk memahami dampak psikologis dan ekonomi dari praktik dark patterns terhadap perilaku konsumen digital. Oleh karena itu, diperlukan metode machine learning yang dapat diterapkan yaitu Support Vector Machine (SVM) untuk mengklasifikasikan polaritas perspektif dari data tidak terstruktur seperti ulasan pada aplikasi X[9].

Selain itu, penggunaan Lexicon InSet dalam proses pelabelan diterapkan pada penelitian ini karena kamus tersebut secara khusus dirancang untuk bahasa Indonesia, sehingga mampu menangkap nuansa sentimen yang lebih sesuai dengan konteks lokal. Dengan demikian, klasifikasi sentimen yang dihasilkan menjadi lebih relevan dan akurat. Pada penelitian tahun 2023, menerapkan algoritma Support Vector Machine (SVM) untuk melakukan klasifikasi dan menunjukkan hasil akurasi yang sangat tinggi, yaitu 98%, dengan nilai precision, recall, dan f1-score yang seimbang (0.98). Studi ini mengonfirmasi efektivitas SVM dalam mengklasifikasikan opini pengguna terhadap aplikasi e-commerce berbasis media sosial[10].

Sementara itu, penelitian lainnya mengembangkan sistem analisis perspektif terhadap aplikasi Tokopedia dengan menggunakan algoritma Support Vector Machine (SVM). Studi ini menekankan pentingnya opini konsumen sebagai indikator kepuasan pengguna, serta menunjukkan bahwa model SVM mampu mengklasifikasikan perspektif dengan akurasi mencapai 81%[11]. Kemudian penelitian pada tahun 2022, algoritma Support Vector Machine (SVM) dimanfaatkan untuk mengklasifikasikan perspektif pengguna terhadap e-commerce Lazada berdasarkan data Twitter. Studi ini mengklasifikasikan perspektif positif, negatif, dan netral dari opini pengguna. Penelitian ini memperkuat bahwa media sosial dapat menjadi sumber data yang valid untuk mengukur kepuasan konsumen terhadap platform e-commerce[12]. Pada tahun 2024, sebuah studi menerapkan analisis perspektif menggunakan metode Support Vector Machine (SVM) terhadap ribuan opini pengguna e-commerce di Twitter.

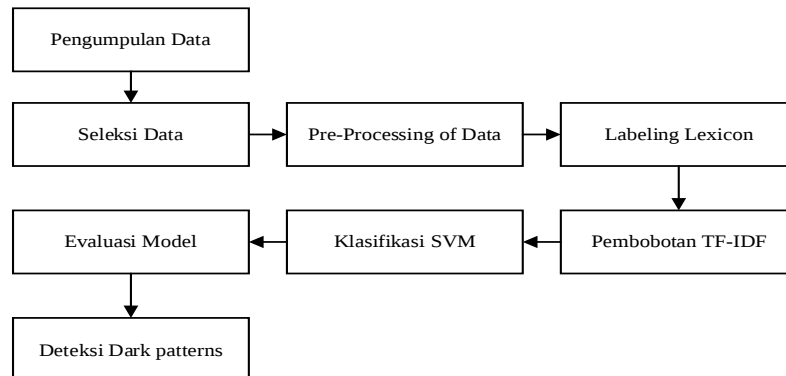
Studi ini berhasil mengidentifikasi bahwa keluhan terkait biaya dan kualitas layanan mendominasi perspektif negatif, serta menegaskan pentingnya media sosial sebagai sumber informasi untuk memahami persepsi pengguna. Penelitian ini tidak secara eksplisit memfokuskan pada platform tertentu seperti Shopee, Tokopedia, atau Lazada[13]. Selain itu, penelitian pada tahun 2021 mengeksplorasi pengaruh dark patterns terhadap tingkat gangguan yang dirasakan pengguna dan kepercayaan terhadap merek dalam konteks belanja online. Hasilnya menunjukkan bahwa meskipun dark patterns seperti Double Negative dan Sneaking dapat meningkatkan konversi, mereka juga meningkatkan rasa terganggu dan menurunkan kepercayaan pengguna terhadap merek. Akan tetapi studi ini menggunakan toko fiktif untuk mengkaji pengalaman pengguna di platform e-commerce[14]. Penelitian ini secara khusus bertujuan untuk menganalisis perspektif konsumen terhadap kebijakan biaya layanan e-commerce dengan fokus pada deteksi indikasi praktik dark patterns di tiga platform terkemuka di Indonesia yaitu Shopee, Tokopedia, dan Lazada.

Pemilihan ketiga platform ini didasarkan pada dominasi pasar dan variasi kebijakan biaya layanan yang diberlakukan, sehingga memungkinkan dilakukannya analisis perspektif konsumen. Menggunakan metode Support Vector Machine (SVM) dan labeling menggunakan Lexicon InSet yaitu kamus sentimen khusus bahasa Indonesia. Melalui pendekatan text mining terhadap data ulasan konsumen, penelitian ini diharapkan dapat mengungkap pola-pola ketidakpuasan konsumen terkait kebijakan e-commerce sekaligus menguji efektivitas metode SVM dalam mendeteksi indikasi dark patterns berdasarkan ekspresi perspektif konsumen di media sosial. Temuan penelitian diharapkan dapat menjadi acuan bagi regulator dan pelaku industri dalam merumuskan kebijakan yang lebih transparan dan melindungi hak-hak konsumen di era perdagangan digital yang semakin kompleks ini.

2. METODOLOGI PENELITIAN

2.1. Alur Penelitian

Berikut ini tahapan penelitian yang dilalui yaitu :



Gambar 1. Tahapan Penelitian

1. Pengumpulan Data

Penelitian ini mengumpulkan ulasan konsumen terkait biaya layanan e-commerce platform X menggunakan teknik web-crawling. Data diambil secara otomatis dengan token autentikasi dari aplikasi X dan library komentar-harvest (Node.js), yang mengekstraksi ulasan berbahasa Indonesia dengan kata kunci seperti "biaya layanan Shopee". Hasilnya disimpan dalam format CSV berisi metadata seperti username, isi ulasan, dan waktu publikasi.

2. Seleksi Data

Proses seleksi data dilakukan untuk memastikan kualitas dan relevansi data yang akan dianalisis. Dari seluruh kolom yang tersedia dalam dataset hasil crawling, penelitian ini hanya memanfaatkan dua variabel utama yaitu username dan konten ulasan (full_text). Variabel-variabel lain yang tidak berkaitan langsung dengan analisis SVM diabaikan untuk menyederhanakan proses penelitian. Selanjutnya dilakukan pembersihan data dengan menghapus ulasan-ulasan konten promosi atau pembahasan yang tidak relevan dengan tujuan penelitian.

3. Pre-Processing of Data

Pre-processing data dilakukan untuk menyiapkan data agar dapat dengan mudah diproses oleh komputer ketika komputer mengklasifikasikan teks menjadi perspektif negatif maupun positif. Tahap dalam preprocessing sebagai berikut.

a. Cleaning

Cleaning yaitu proses pembersihan ulasan dengan tujuan untuk menghapus tanda baca, mention, tagar, link, dan karakter lainnya pada dokumen.

b. Case Folding

Case folding yaitu proses menyamakan karakter dari setiap huruf dalam kalimat atau kata dalam dataset untuk dikonversi menjadi huruf kecil atau lowercase.

c. Tokenizing

Tokenizing digunakan untuk membagi teks atau kalimat menjadi unit-unit yang lebih kecil yang disebut token, biasanya berupa kata, frasa, atau simbol.

d. Stopword Removing

Stopwords removal adalah proses yang dilakukan untuk menghilangkan kata-kata yang tidak memiliki makna. Tahap ini akan menggunakan kamus stopwords. Kata-kata yang terdapat dalam daftar stopwords akan dihapus.

e. Stemming

Stemming yaitu proses yang dilakukan untuk mengubah kata-kata yang tidak baku menjadi baku dengan menghilangkan imbuhan awalan dan akhiran.

4. Labeling Lexicon

Pelabelan digunakan untuk mengelompokkan ulasan ke dalam kategori positif dan negatif berdasarkan batasan atau kriteria tertentu. Pada penelitian ini dilakukan dengan bantuan kamus Lexicon InSet, yaitu kamus perspektif bahasa Indonesia yang tersedia di repository GitHub. Kamus tersebut membantu dalam memberikan skor pada setiap kata yang ada pada dataset. Lexicon InSet merupakan basis data kosakata untuk analisis perspektif dalam bahasa Indonesia yang berisi 3.609 entri kata positif dan 6.609 kata negatif, dikembangkan berdasarkan data dari platform X. Setiap kata diberi nilai skala manual dari -5 hingga +5 untuk penilaian perspektif[15].

5. Pembobotan TF-IDF

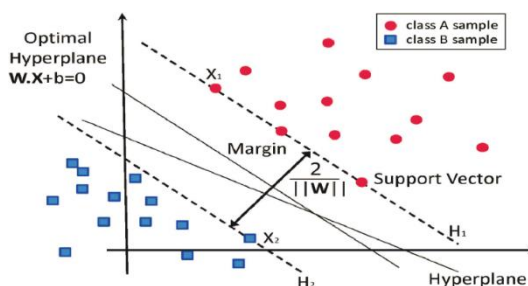
Metode TF-IDF adalah teknik yang digunakan untuk menghitung bobot setiap kata dalam proses pengambilan informasi. Metode ini populer karena efisiensi, kemudahan penggunaan, dan keakuratan hasilnya. Prinsip

utama TF-IDF adalah memberikan nilai bobot berdasarkan hubungan antara suatu kata (term) dengan suatu dokumen. TF-IDF berfungsi sebagai indikator statistik untuk menilai seberapa penting suatu kata dalam sebuah dokumen atau kumpulan dokumen. Dalam penerapannya pada dokumen tunggal, setiap kalimat dapat diperlakukan sebagai dokumen terpisah. Frekuensi kemunculan kata dalam suatu dokumen menunjukkan tingkat relevansinya. Bobot suatu kata akan semakin tinggi jika kata tersebut sering muncul dalam dokumen tersebut, tetapi akan lebih rendah jika kata tersebut juga muncul di banyak dokumen lain[16]. Metode ini menghitung bobot kata dalam sebuah dokumen dilakukan dengan mengalikan nilai Term Frequency dengan Inverse Document Frequency[17]. Rumus untuk menghitung TF-IDF diberikan pada persamaan (1):

$$W_{dt} = t f_{dt} \times i d f_{dt} = t f_{dt} \times \log \frac{D}{df} \quad (1)$$

6. Klasifikasi SVM

Support Vector Machine (SVM) merupakan algoritma klasifikasi yang pertama kali dikembangkan oleh Boser, Guyon, dan Vapnik pada tahun 1992. SVM merupakan algoritma machine learning yang belajar dari data yang sudah berlabel dan berdasarkan prinsip dasar teori statistik dan memiliki cakupan yang sangat luas, seperti pendeteksian intrusi, klasifikasi ekspresi wajah, prediksi data deret waktu, pengenalan suara, pengenalan gambar, pengolahan sinyal, deteksi gen, klasifikasi teks, pengenalan jenis huruf, diagnosis kerusakan, analisis kimia, dan bidang lainnya[18]. Teknik ini banyak dipilih karena tingkat akurasi yang tinggi serta kemampuannya dalam menangani data dengan dimensi besar[19]. Support Vector Machine merupakan salah satu algoritma machine learning yang dirancang untuk menemukan hyperplane optimal guna memisahkan dua kelas berbeda. SVM beroperasi dengan membagi data ke dalam dua kategori, yaitu kelas 1 dan kelas -1. Sebagai contoh, pada Gambar 2, data berlabel -1 direpresentasikan dengan titik berwarna biru, sedangkan data kelas 1 ditandai dengan warna merah. Solusi permasalahan ini diperoleh dengan menentukan sebuah batas pemisah (hyperplane) yang optimal untuk memisahkan kedua kelas tersebut[20].



Gambar 2. Ilustrasi Hyperline SVM[21].

Pada Gambar 2 di atas menunjukkan hyperline yang memisahkan 2 kelas yang berbeda secara linier, garis putus-putus sebagai margin atau jarak antar hyperline dengan titik yang dilingkari (pattern) terdekat yang merupakan SVM[22]. Dengan pendekatan ini memungkinkan SVM memiliki performa generalisasi yang unggul dalam menangani data yang belum pernah dilihat sebelumnya[16].

7. Evaluasi Model

Penelitian ini menggunakan Confusion Matrix sebagai dasar untuk mengevaluasi performa sistem klasifikasi yang dikembangkan. Confusion matriks yang mencakup accuracy (akurasi), precision (presisi), recall, dan f1-score[23]. Rumus-rumus perhitungan untuk masing-masing matrik evaluasi tersebut akan diuraikan secara rinci dalam persamaan di bawah yang mengacu pada literatur[24].

Accuracy: Mengukur seberapa banyak prediksi yang benar dari total data.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Precision : Mengukur seberapa akurat model dalam memprediksi positif. Meminimalkan false positif.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Recall : Mengukur seberapa baik model menemukan semua kasus positif yang benar.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F1-score : Rata-rata harmonik dari precision dan recall. Cocok untuk dataset yang tidak seimbang, karena menyeimbangkan keduanya.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

8. Deteksi Dark patterns

Pendeteksian dilakukan dengan mendeteksi indikasi keberadaan dark patterns pada komentar yang telah diklasifikasikan sebagai sentimen negatif. Proses deteksi ini dilakukan melalui pendekatan berbasis kata kunci, yang dikembangkan dari referensi literatur terkait desain manipulatif dalam antarmuka digital. Kategori dark patterns yang dianalisis dalam penelitian ini mencakup Hidden Costs, Sneak into Basket, Obstruction, dan Forced Action. Deteksi dilakukan sebagai tahap lanjutan untuk mengidentifikasi pola-pola interaksi yang berpotensi merugikan pengguna berdasarkan ekspresi ketidakpuasan dalam komentar yang telah dikumpulkan.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Proses pengumpulan data menggunakan tweet-harvest yaitu alat berbasis node.js yang membantu dalam proses crawling data dan token auth login serta memasang limit ulasan sebesar 500, sehingga proses crawling dapat menghasilkan banyak ulasan dan memudahkan dalam proses analisis. Proses ini dilakukan dengan membuat batasan waktu dari bulan Oktober tahun 2022 hingga Maret 2025. Kata kunci yang digunakan ialah “biaya layanan shopee”, “biaya layanan tokped” dan “biaya layanan lazada”. Hasil crawling dapat diamati pada tabel 1.

Tabel 1. Hasil *Crawling* Ulasan pada Aplikasi X

Nama Aplikasi	Jumlah Dataset
Shopee	503 kalimat opini
Tokopedia	385 kalimat opini
Lazada	427 kalimat opini

3.2 Seleksi Data

Seleksi data dilakukan dengan menggunakan library pandas dan mengambil beberapa keyword yang relevan untuk ulasan yang akan dipertahankan, kemudian penghapusan pada ulasan dari akun resmi e-commerce agar mempertahankan kualitas review dari pengguna. Hasil seleksi data menunjukkan pengurangan jumlah data setelah proses dilakukan pada setiap dataset aplikasi, dimana data Shopee berkurang dari 503 menjadi 427 data, Lazada dari 427 menjadi 420 data, kecuali Tokopedia dimana jumlah data tetap sebesar 385.

Tabel 2. Contoh Dataset yang Relevan dan Tidak Relevan pada Aplikasi Shopee

No	Username	Ulasan	Keterangan
1.	w1nc0metru3	Apasih shopee biaya layanan lu gede bgt 8k hampir 10k gila	Relevan
2.	@ShopeeID	@sebutakuno Hi kak sebutakuno mohon maaf untuk kendalanya aku bantu infoin ya kak mohon pastikan saldo ShopeePay yang Kakak miliki cukup atau lebih dari subtotal jumlah produk ongkos kirim dan biaya layanan ya	Tidak Relevan

Pada Tabel 2 ulasan dari username w1nc0metru3 ini berasal dari pengguna biasa, bukan akun resmi Shopee. Isi ulasan mengandung keluhan secara eksplisit terhadap biaya layanan yang dikenakan oleh Shopee. Kata kunci seperti “biaya layanan” dan “shopee” disebutkan secara langsung, dan opini pengguna sangat jelas dimana mengungkapkan kekecewaan karena nominal biaya yang dianggap tinggi. Oleh karena itu, ulasan ini dikategorikan relevan, tidak seperti contoh kedua karena ulasan tersebut berasal dari akun resmi Shopee sehingga tidak dapat dikategorikan sebagai ulasan relevan.

3.3 Pre-Processing Data

Data yang telah diseleksi selanjutnya melalui tahap pre-processing yang meliputi pembersihan data (cleaning), konversi huruf menjadi huruf kecil (case folding), tokenisasi (tokenizing), penghapusan kata-kata umum (stopword removal), serta stemming. Pada tahap ini, digunakan pendekatan fuzzy string matching untuk membantu proses normalisasi kata, khususnya dalam menangani kata-kata slang atau tidak baku yang umum dijumpai dalam teks informal. Fuzzy string matching memungkinkan sistem mengenali kemiripan antar kata meskipun terdapat kesalahan penulisan (typo), variasi bentuk, atau perbedaan minor lainnya. Dalam penelitian ini, digunakan library FuzzyWuzzy yang memanfaatkan algoritma Levenshtein Distance untuk mengukur tingkat kemiripan antar string. Pendekatan ini digunakan untuk mencocokkan kata slang dalam teks dengan padanan katanya dalam kamus slang yang telah disusun sebelumnya. Jika kemiripan antara kata dalam teks dan entri dalam kamus melebihi ambang batas tertentu (dalam penelitian diatur sebesar 90), maka kata tersebut akan digantikan dengan bentuk bakunya.

Dengan cara ini, proses normalisasi dapat dilakukan secara lebih fleksibel dan toleran terhadap variasi penulisan, sehingga sangat sesuai untuk pengolahan data teks berbahasa informal seperti ulasan di media sosial X.

Tabel 3. Kamus Bahasa Slang

Kata Tidak Baku	Kata Baku
"yg"	"yang"
"ga"	"tidak"
"dah"	"sudah"
"co"	"checkout"
"ga tau"	"tidak tahu"
"byr"	"bayar"

Pada Tabel 3 merupakan contoh dari isi kamus bahasa slang yang dibuat secara manual menyesuaikan dataset yang ada, seperti kata "yg" memiliki kata baku "yang". Proses ini akan memudahkan pada proses stopwords removing dimana kata yang termasuk pada stopwords akan dihapus seperti kata "yang". Jika tidak dilakukan proses penggantian kata dengan kamus slang maka kata-kata tersebut akan tetap muncul sehingga sulit dipahami dan dianggap sebagai kata baku oleh sistem.

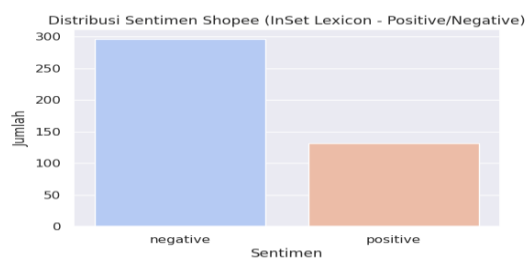
Proses pre-processing ini menghasilkan korpus teks yang telah terstandarisasi dan siap untuk dilakukan analisis lebih lanjut, sekaligus mempertahankan makna asli dari ulasan pengguna meskipun telah dinormalisasi ke dalam bentuk yang lebih baku. Pada tabel 4 merupakan contoh hasil dari pre-processing data dimana ulasan nomor 2 misalnya terdapat kata slang "bgt" yang kemudian diubah sistem menjadi "banget" dan tidak terhapus karena tidak terdeteksi dalam kamus stopwords.

Tabel 4. Hasil Pre-Processing Dataset Aplikasi Shopee

No	Username	Ulasan	CleanedReview
1.	50AR_	@kimmyexp Lah.. biaya layanan shopee naek jd 1900 https://t.co/O9HVPbbf0	biaya layan shopee naek
2.	w1nc0metru3	Apasih shopee biaya layanan lu gede bgt 8k hampir 10k gila	shopee biaya layan lu gede banget gila
3.	zavvac	Biaya layanan shopee mahal bgtttttt	biaya layan shopee mahal banget

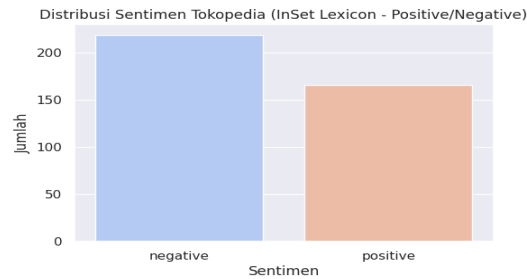
3.4 Pelabelan Data

Pelabelan dilakukan menggunakan metode lexicon-based dengan bantuan Lexicon InSet dari sumber repository Github. Program ini menganalisis perspektif ulasan pengguna dengan membandingkan kata-kata dalam teks terhadap kamus kata positif dan negatif dari Lexicon InSet. Pertama, kamus perspektif dibaca dan disimpan dalam bentuk set untuk pencarian yang lebih cepat. Data ulasan pengguna yang sudah diproses kemudian dimuat untuk dianalisis. Setiap ulasan diperiksa kata per kata. Jika kata positif lebih banyak, ulasan dianggap positif; jika kata negatif lebih banyak atau sama, ulasan dianggap negatif. Hasilnya ditampilkan dalam visualisasi grafik batang, dengan warna biru untuk yang jumlah datanya lebih banyak dan merah untuk yang jumlah datanya lebih sedikit seperti pada gambar 3 hingga 5.



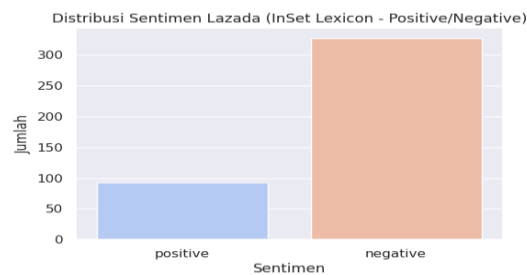
Gambar 3. Hasil Pelabelan Aplikasi Shopee

Pada Gambar 3 hasil pelabelan menggunakan Lexicon InSet pada ulasan Shopee didominasi oleh kategori negatif dengan 296 data (69,3%), sementara kategori positif hanya mencakup 131 data (30,7%).



Gambar 4. Hasil Pelabelan Aplikasi Tokopedia

Pada Gambar 4 hasil pelabelan menggunakan Lexicon InSet pada ulasan Tokopedia juga didominasi oleh kategori negatif dengan 219 data (56,9%), sementara kategori positif hanya mencakup 166 data (43,1%).



Gambar 5. Hasil Pelabelan Aplikasi Lazada

Pada Gambar 5 hasil pelabelan menggunakan Lexicon InSet pada aplikasi Lazada menghasilkan distribusi kategori negatif sebesar 327 data (77,86%) dan positif sebesar 93 data (22,14%).

3.5 Pembobotan TF-IDF

Pembobotan dilakukan untuk melihat seberapa penting kata tersebut dalam NLP dimana semakin sering muncul kata dalam satu dokumen maupun dalam seluruh dokumen maka semakin kecil angkanya untuk diperhitungkan.

Tabel 5. Hasil Pembobotan TF-IDF Aplikasi Shopee

Term	TF	IDF	TF-IDF
biaya	425	1.005865	0.152328
layan	424	1.005865	0.138968
shopee	422	1.008811	0.136317
banget	69	2.723937	0.041698
checkout	63	2.922788	0.039648
mahal	42	3.197225	0.039645
tangan	37	3.467515	0.027635
admin	47	3.145931	0.026592
bayar	48	3.097141	0.025436
kalo	40	3.338303	0.025085

Pada Tabel 5, kata “biaya” menjadi kata paling penting dengan jumlah TF-IDF terbesar dengan skor 0.152328 dan diikuti oleh term “layan” yang merujuk pada frasa biaya layanan. Ini menandakan bahwa para pengguna mengeluhkan soal biaya layanan yang diberlakukan oleh aplikasi Shopee.

Tabel 6. Hasil Pembobotan TF-IDF Aplikasi Tokopedia

Term	TF	IDF	TF-IDF
biaya	384	1.000000	0.119036
tokped	381	1.009756	0.099072
layan	381	1.013030	0.098772
aplikasi	66	2.690290	0.041950
kak	41	3.206981	0.041670
dapet	45	3.069780	0.041392
incl	29	3.555287	0.036357
checkout	54	2.904700	0.035999

jasa	47	3.044462	0.033123
banget	58	2.862140	0.032752

Pada Tabel 6 kata “biaya” menjadi kata paling penting dengan jumlah TF-IDF terbesar dengan skor sebesar 0.119036 dan diikuti term “tokped”. Ini bisa jadi menandakan bahwa para pengguna mengeluhkan soal biaya layanan yang diberlakukan oleh aplikasi Tokopedia.

Tabel 7. Hasil Pembobotan TF-IDF Aplikasi Lazada

Term	TF	IDF	TF-IDF
biaya	420	1.000000	0.146645
lazada	403	1.045531	0.140564
layan	419	1.002972	0.133287
kak	104	2.377432	0.088491
belanja	117	2.235115	0.070941
checkout	104	2.389266	0.067604
pake	57	2.928263	0.046107
kalo	50	3.058883	0.039127
banget	53	2.991442	0.036447
admin	37	3.354347	0.028780

Pada Tabel 7 kata “biaya” menjadi kata paling penting dengan jumlah TF-IDF terbesar dengan skor 0.146645 dan diikuti oleh term “lazada”. Ini juga menandakan bahwa para pengguna mengeluhkan soal biaya layanan yang diberlakukan oleh aplikasi Lazada.

3.6 Analisis Perspektif Konsumen Menggunakan SVM

Proses analisis perspektif pengguna menggunakan metode SVM. Pertama perlu dilakukan proses split data yang sudah diberi label dengan pembagian 80 data training 20 testing. Berdasarkan hasil tuning hyperparameter dengan metode RandomizedSearchCV untuk mencari nilai parameter terbaik dengan diatur iterasi maksimal 50 kali. Pada model Support Vector Machine (SVM) terbaik diperoleh aplikasi shopee dengan penggunaan kernel linear serta nilai parameter C sebesar 100.0 dan gamma 0.01. Nilai akurasi dari hasil training sangat tinggi yaitu 1.0 yang menandakan sistem berhasil mendeteksi positif dan negatif dengan baik. Konfigurasi ini juga dilakukan pada model untuk aplikasi Lazada, yang menunjukkan performa optimal pada nilai C yang berbeda yaitu 1.0 dan nilai akurasi training sebesar 1.0. Sementara itu, untuk aplikasi Tokopedia, nilai C yang didapatkan yaitu 10.0 dengan akurasi training sebesar 0.99. Pemilihan kernel linear mengindikasikan bahwa data perspektif yang digunakan dapat dipisahkan secara efektif menggunakan garis lurus di ruang fitur. Setelah proses pelatihan selesai dilakukan, tahap selanjutnya yaitu evaluasi. Hasil evaluasi metode SVM pada ketiga aplikasi ditampilkan dalam tabel berikut:

Tabel 8. Evaluasi SVM Aplikasi Shopee

	Precision	Recall	F1-Score	Support
negative	0.85	0.87	0.86	60
positive	0.68	0.65	0.67	26
accuracy			0.80	86
macro avg	0.77	0.76	0.76	86
weighted avg	0.80	0.80	0.80	86

Pada Tabel 8 model untuk Shopee berhasil mencapai akurasi tinggi sebesar 80%, dengan performa cukup baik pada perspektif negatif dengan F1-score 0.86. Pada perspektif positif, model memiliki precision yang cukup rendah yaitu 0.68, artinya setiap prediksi positif benar, namun recall 0.65 menunjukkan bahwa masih ada ruang untuk meningkatkan deteksi terhadap perspektif positif.

Tabel 9. Evaluasi SVM Aplikasi Tokopedia

	Precision	Recall	F1-Score	Support
negative	0.76	0.84	0.80	44
positive	0.75	0.64	0.69	33
accuracy			0.75	77
macro avg	0.75	0.74	0.74	77
weighted avg	0.75	0.75	0.75	77

Pada Tabel 9 model Tokopedia juga mencatat akurasi 75%, dengan F1-score 0.80 untuk perspektif negatif, menunjukkan keseimbangan yang sangat baik antara precision dan recall. Untuk perspektif positif, precision 0.75 menunjukkan bahwa setiap prediksi positif adalah benar, namun recall 0.64 mengindikasikan bahwa hanya sebagian perspektif positif yang terdeteksi.

Tabel 10. Evaluasi SVM Aplikasi Lazada

	Precision	Recall	F1-Score	Support
negative	0.88	0.97	0.92	65
positive	0.83	0.53	0.65	19
accuracy			0.87	84
macro avg	0.85	0.75	0.78	84
weighted avg	0.87	0.87	0.86	84

Pada Tabel 10 model Lazada mencatat akurasi tertinggi yaitu 87% dengan presisi 0.88 dan recall 0.97 yang artinya model dapat mengenali teks negatif dengan baik. Berbeda dengan perspektif positif yang memiliki presisi 1.00 yang menandakan sistem bekerja dengan baik dalam mengenali kata positif tapi memiliki recall 0.33, menunjukkan bahwa deteksi terhadap perspektif positif masih terbatas. Hal ini dikarenakan ketidakseimbangan jumlah data positif yang tersedia.

3.7 Deteksi Dark Patterns

Dalam penelitian ini, proses deteksi dark pattern dilakukan melalui pendekatan berbasis kata kunci yang dikembangkan secara tematik dari hasil observasi ulasan pengguna. Setiap ulasan dianalisis berdasarkan keberadaan frasa atau ekspresi yang memiliki kemiripan makna dengan karakteristik dark pattern sebagaimana dikemukakan dalam penelitian tahun 2019[6]. Untuk mengakomodasi variasi bahasa informal di media sosial, digunakan metode pencocokan fuzzy string matching, yang memungkinkan sistem mendeteksi pola meskipun ekspresi pengguna tidak identik secara literal dengan kata kunci yang telah ditentukan. Deteksi ini mengacu pada empat kategori dark pattern utama yang relevan dengan fokus penelitian pada biaya layanan e-commerce, yaitu hidden costs, sneak into basket, obstruction, dan forced action. Ulasan-ulasan yang telah diberi label perspektif negatif dianalisis untuk mengidentifikasi indikasi keberadaan dark pattern melalui kata kunci tersebut. Hasil deteksi ini disajikan dalam bentuk tabel ringkasan yang menampilkan contoh representatif dari masing-masing kategori berdasarkan tiga platform e-commerce yang diteliti, yaitu Shopee, Tokopedia, dan Lazada.

Tabel 11. Hasil Identifikasi Indikasi Dark Patterns pada Aplikasi Shopee

Username	Ulasan	Kategori
fluffyblue	Bjir lah shopee lu kayak ngerampok dah. Gak daftar xtra tau tau otomatis jadi xtra trus kena biaya layanan https://t.co/dHodH7ujnw	<i>Sneak into basket</i>
mahardixxx	@ShopeeID katanya pro UMKM biaya layanan tinggi banget terus berhenti Dari program Cashback Xtra & gratis ongkir Xtra susah banget' di paksa join program apa?! Makin gajelas shopee.. pindah tiktokshop aja gaesss	<i>Obstruction</i>
itsmedeev_	baru tau shopee ada biaya layanan 10k najiss	<i>Hidden cost</i>

Pada Tabel 11 di atas teridentifikasi memiliki indikasi dark patterns dengan ulasan yang merujuk pada ketidaktahuan seorang pengguna pada kebijakan yang diberlakukan oleh aplikasi Shopee. Identifikasi pertama adanya indikasi dark patterns kategori sneak into basket dimana mengindikasikan bahwa pengguna dikenakan biaya layanan tambahan yang tidak pernah didaftarkan atau setuju secara eksplisit. Hal ini sesuai dengan konsep sneak into basket yaitu strategi di mana suatu biaya atau layanan ditambahkan secara otomatis ke dalam transaksi tanpa konfirmasi eksplisit dari pengguna. Praktik ini sering digunakan untuk mengelabui pengguna agar membayar lebih tanpa disadari, dan dapat dianggap sebagai bentuk penyisipan biaya tersembunyi yang manipulatif. Pada ulasan kedua, tampak bahwa pengguna mengalami kebingungan akibat ketidakkonsistenan biaya layanan antar akun pengguna Shopee. Ketika sistem menetapkan biaya yang berbeda tanpa penjelasan yang transparan, hal ini dapat menghambat pemahaman pengguna dalam mengambil keputusan rasional dalam bertransaksi. Ketika informasi biaya tidak disampaikan secara konsisten, pengguna menjadi tidak memiliki kontrol penuh atas proses pengambilan keputusan finansialnya. Pada ulasan ketiga menunjukkan bahwa pengguna baru menyadari

keberadaan biaya layanan setelah berada di tahap akhir transaksi. Tidak adanya informasi sejak awal tentang biaya tambahan tersebut menjadikan pengalaman pengguna terasa mengejutkan dan merugikan. Hidden cost adalah salah satu dark pattern paling umum ditemukan di e-commerce, dan bertujuan untuk menyembunyikan total biaya hingga pengguna berada pada titik di mana mereka cenderung tidak membatalkan pembelian.

Tabel 12. Hasil Identifikasi Indikasi Dark Patterns pada Aplikasi Tokopedia

Username	Ulasan	Kategori
crxbbx	@convomf Jgnkan shopee tokped pun gini biaya app biaya layanan sama tbtt ada wajib asuransi pengiriman padahal ini barangnya mainan doang https://t.co/8pnp2vrooI	<i>Sneak into basket</i>
vntcel	Aih dah lama gapernah co di tokped tbtt kaget dah ada biaya jasa aplikasinya mana 2rb udah nambah biaya layanan kerasa berat ini ditambah pula ini hiks https://t.co/QaXQp02zcX	<i>Hidden cost</i>

Sama halnya dengan aplikasi shopee, pada dataset hasil analisis ulasan Tokopedia, di temukan indikasi 2 jenis dark patterns, yaitu sneak into basket dan hidden cost. Ulasan pertama menyebutkan bahwa biaya asuransi ditambahkan secara otomatis dan wajib dibayarkan tanpa persetujuan oleh pengguna. Sedangkan untuk ulasan kedua dimana menunjukkan ekspresi terkejut karena adanya biaya tambahan yaitu biaya jasa aplikasi yang baru diketahui di akhir sehingga biaya tersebut dirasa membebani si pengguna.

Tabel 13. Hasil Identifikasi Indikasi Dark Patterns pada Aplikasi Tokopedia

Username	Ulasan	Kategori
Christineyuwan	Loh lazada sekarang kena biaya layanan jg	<i>Hidden cost</i>

Pada tabel 13 menunjukkan indikasi dark patterns pada biaya layanan lazada hanya 1 jenis yaitu hidden cost, dimana pengguna merasa terkejut mendapati biaya layanan yang dikenakan ketika membeli barang di platform tersebut. Ini didukung dengan banyaknya ulasan yang berisi informasi bahwa lazada sebelumnya tidak memberlakukan biaya layanan, akan tetapi terdapat kebijakan baru sehingga terdapat biaya layanan dan indikasi bahwa kebijakan kurang transparan sehingga pengguna tidak mengetahui kebijakan tersebut.

4. KESIMPULAN

Berdasarkan hasil analisis perspektif konsumen terhadap tiga platform e-commerce (Shopee, Tokopedia, dan Lazada) menggunakan metode lexicon-based dan Support Vector Machine (SVM), dapat disimpulkan bahwa mayoritas ulasan pengguna didominasi oleh sentimen negatif. Shopee mencatat 69,3% ulasan negatif, Tokopedia 56,9%, dan Lazada bahkan mencapai 77,86%. Hal ini menunjukkan bahwa pengguna cenderung lebih vokal dalam menyampaikan ketidakpuasan, terutama terkait biaya layanan yang dianggap tidak transparan dan terlalu tinggi. Analisis TF-IDF menguatkan temuan ini, dengan kata kunci seperti "biaya", "layan", dan "mahal" mendominasi bobot tertinggi di ketiga platform, mengindikasikan bahwa isu biaya tambahan menjadi keluhan utama. Metode SVM menunjukkan akurasi yang cukup baik dalam mengklasifikasikan sentimen, dengan nilai akurasi tertinggi pada Lazada (87%), diikuti Shopee (80%) dan Tokopedia (75%). Namun, model cenderung lebih efektif mendeteksi ulasan negatif dibandingkan positif, yang tercermin dari recall dan precision yang lebih tinggi untuk kategori negatif. Ketimpangan ini dapat disebabkan karena perbedaan jumlah data positif yang sedikit dan juga akibat variasi ekspresi positif yang lebih beragam serta penggunaan bahasa informal atau slang, sementara keluhan negatif cenderung menggunakan kata-kata spesifik dan repetitif seperti "mahal". Lebih lanjut, identifikasi dark patterns mengungkap praktik manipulatif yang kerap dikeluhkan pengguna, seperti hidden cost (biaya tersembunyi), sneak into basket (biaya otomatis tanpa konfirmasi), dan obstruction (sistem yang membingungkan). Shopee dan Tokopedia menunjukkan indikasi kuat terhadap ketiga pola tersebut, sementara Lazada terutama terkait hidden cost. Ulasan seperti "tau-tau kena biaya" atau "tiba-tiba wajib asuransi" mencerminkan kurangnya transparansi dan etika bisnis, yang berpotensi merusak kepercayaan konsumen. Secara keseluruhan, penelitian ini menyoroti perlunya perbaikan dalam transparansi kebijakan biaya, desain antarmuka yang lebih jujur, serta peningkatan layanan pelanggan. Temuan ini juga menjadi masukan bagi regulator untuk mengawasi praktik dark patterns di e-commerce, demi menciptakan ekosistem digital yang lebih adil dan berorientasi pada kepuasan pengguna.

REFERENCES

- [1] A. Utami, M. Irwan, and P. Nasution, "Perkembangan Pasar Online (E-Commerce) Di Era Modern Dan Pengaruhnya

- Terhadap Kepercayaan Konsumen,” *J. Ekon. Manaj. Dan Bisnis*, vol. 1, no. 2, pp. 126–132, 2023, [Online]. Available: <https://doi.org/XX.XXXXX/JMEB>
- [2] R. Godara, “E-Commerce Trends Post Covid-19,” *Shodh Sarita*, vol. 8, no. 29, pp. 187–191, 2022.
 - [3] A. Wardhana, *E-Commerce In The Digital Edge – Edisi Indonesia*, Indonesia., no. September. Purbalingga: CV. Eureka Media Aksara, 2024.
 - [4] O. Thoyib and H. Wijaya, “E-Commerce: Perkembangan, Tren, dan Peraturan Perundang-Undangan,” *J. Ilm. Ekon. DAN BISNIS*, vol. 16, no. 1, pp. 41–47, 2023, doi: 10.51903/e-bisnis.v16i1.
 - [5] C. M. Gray, Y. Kou, B. Battles, J. Hoggatt, and A. L. Toombs, “The dark (patterns) side of UX design,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–14, 2018, doi: 10.1145/3173574.3174108.
 - [6] A. Mathur *et al.*, “Dark patterns at scale: Findings from a crawl of 11K shopping websites,” *Proc. ACM Human-Computer Interact.*, vol. 3, no. CSCW, p. art. no. 81, 2019, doi: 10.1145/3359183.
 - [7] A. Mathur and J. Mayer, “What makes a dark pattern... dark? design attributes, normative considerations, and measurement methods,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–18, 2021, doi: 10.1145/3411764.3445610.
 - [8] W. C. Koh and Y. Z. Seah, “Unintended consumption: The effects of four e-commerce dark patterns,” *Clean. Responsible Consum.*, vol. 11, no. September 2022, p. 100145, 2023, doi: 10.1016/j.clrc.2023.100145.
 - [9] H. Syahputra, “Sentiment Analysis of Community Opinion on Online Store in Indonesia on Twitter using Support Vector Machine Algorithm (SVM),” *J. Phys. Conf. Ser.*, vol. 1819, no. 1, 2021, doi: 10.1088/1742-6596/1819/1/012030.
 - [10] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, “Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM),” *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023, doi: 10.37905/jjee.v5i1.16830.
 - [11] D. Fadila and M. Ikhsan, “Analisis Sentimen Pada Aplikasi Tokopedia Menggunakan Metode Support Vector Machine,” *Progresif Jurnal Ilm. Komput.*, vol. 21, no. 1, pp. 238–247, 2025, doi: 10.35889/progresif.v21i1.2593.
 - [12] A. J. Putri, A. S. Syafira, M. E. Purbaya, and D. Purnomo, “Analisis Sentimen E-Commerce Lazada pada Jejaring Sosial Twitter Menggunakan Algoritma Support Vector Machine,” *J. TRINISTIK J. Tek. Ind. Bisnis Digit. dan Tek. Logistik*, vol. 1, no. 1, pp. 16–21, 2022, doi: 10.20895/trinistik.v1i1.447.
 - [13] M. Hamka and Tukiran, “Analisis Sentimen Pengguna E-Commerce dan Marketplace Menggunakan Support Vector Machine,” *J. Rekamaya Sist. Inf. dan Teknol.*, vol. 1, no. 4, pp. 273–282, 2024, doi: 10.59407/jrsit.v1i4.555.
 - [14] C. Voigt, S. Schlögl, and A. Groth, “Dark Patterns in Online Shopping: of Sneaky Tricks, Perceived Annoyance and Respective Brand Trust,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12783 LNCS, pp. 143–155, 2021, doi: 10.1007/978-3-030-77750-0_10.
 - [15] F. Koto and G. Y. Rahmaningtyas, “Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs,” *2017 Int. Conf. Asian Lang. Process.*, pp. 391–394, 2017, doi: 10.1109/IALP.2017.8300625.
 - [16] R. Ramlan, N. Satyahadewi, and W. Andani, “Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM,” *Jambura J. Math.*, vol. 5, no. 2, pp. 431–445, 2023, doi: 10.34312/jjom.v5i2.20860.
 - [17] D. Robison Manalu, M. Christofell, L. Tobing, and M. Yohanna, “METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi ANALISIS SENTIMEN TWITTER TERHADAP WACANA PENUNDAAN PEMILU DENGAN METODE SUPPORT VECTOR MACHINE,” *Menthomika*, vol. 6, no. 2, pp. 149–156, 2022, [Online]. Available: <https://doi.org/10.46880/jmika.Vol6No2.pp149-156>
 - [18] D. Mustafa Abdullah and A. Mohsin Abdulazeez, “Machine Learning Applications based on SVM Classification A Review,” *Qubahan Acad. J.*, vol. 1, no. 2, pp. 81–90, Apr. 2021, doi: 10.48161/qaj.v1n2a50.
 - [19] Y. I. Kurniawan, A. Rahmawati, N. Chasanah, and A. Hanifa, “Application for determining the modality preference of student learning,” *J. Phys. Conf. Ser.*, vol. 1367, no. 1, 2019, doi: 10.1088/1742-6596/1367/1/012011.
 - [20] Y. Guo, S. Han, Y. Li, C. Zhang, and Y. Bai, “K-Nearest Neighbor combined with guided filter for hyperspectral image classification,” *Procedia Comput. Sci.*, vol. 129, pp. 159–165, 2018, doi: 10.1016/j.procs.2018.03.066.
 - [21] I. Setiawan *et al.*, “Credit Risk Management Prediction Using the Support Vector Machine (SVM) Algorithm,” *Proc. 1st Int. Conf. Mech. Ind. Mater. Eng. (MIMSE-I-C 2022)*, vol. 102, pp. 195–206, 2022, doi: 10.2991/978-94-6463-084-8_18.
 - [22] I. MAHENDRO and D. ABIMANTO, *ALGORITMA SUPPORT VECTOR MACHINE UNTUK MENGANALISA KEPUASAN MAHASISWA TERHADAP PEMBELAJARAN E-LEARNING*. Semarang: Cv Pustaka Stimart AMINI, 2022.
 - [23] M. H. Aufan, M. R. Handayani, A. B. Nurjanna, and N. C. Hendro, “THE PERCEPTIONS OF SEMARANG FIVE STAR HOTEL TOURISTS WITH SUPPORT VECTOR MACHINE ON GOOGLE REVIEWS PERSEPSI WISATAWAN HOTEL BINTANG LIMA SEMARANG DENGAN,” vol. 5, no. 5, pp. 1241–1247, 2025.
 - [24] M. R. Zamroni, R. A. Hamid, S. Mujilahwati, and M. Sholihin, “DETECTION OF BULLYING CONTENT IN ONLINE NEWS USING A COMBINATION OF RoBERTa-BiLSTM DETEKSI KONTEN PERUNDUNGAN DALAM BERITA ONLINE MENGGUNAKAN KOMBINASI RoBERTa-BiLSTM,” vol. 6, no. 1, pp. 41–50, 2025.