

# ANALISIS KEAMANAN SISTEM INFORMASI MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE TERHADAP PENGGUNA SHOPEE

(INFORMATION SYSTEM SECURITY ANALYSIS USING SUPPORT VECTOR MACHINE ALGORITHM ON SHOPEE USERS)

Muhamad Dody Firmansyah<sup>1</sup>, Christopher<sup>2</sup>, Mangapul Siahaan<sup>3</sup>

<sup>1,2,3,4</sup>Sistem Informasi Universitas Internasional Batam

e-mail: [dody.firmansyah@uib.edu](mailto:dody.firmansyah@uib.edu)<sup>1</sup>, [2231129.christopher@uib.edu](mailto:2231129.christopher@uib.edu)<sup>2</sup>, [mangapul.siahaan@uib.ac.id](mailto:mangapul.siahaan@uib.ac.id)<sup>3</sup>

## ABSTRAK

Perkembangan e-commerce di Indonesia sangat bergantung pada keamanan sistem informasi, khususnya pada platform Shopee, yang memiliki banyak transaksi. Peningkatan frekuensi penggunaan aplikasi dapat meningkatkan potensi bahaya seperti penyalahgunaan akun, akses tidak sah, dan aktivitas mencurigakan. Oleh karena itu, analisis yang didasarkan pada karakteristik dan tindakan pengguna diperlukan untuk memahami risiko keamanan. Metode pembelajaran mesin digunakan dalam penelitian ini untuk mengevaluasi pola risiko keamanan pengguna Shopee. Variabel demografi, frekuensi penggunaan, tingkat kewaspadaan, dan pengalaman dengan aktivitas mencurigakan dihitung dalam kuesioner yang diberikan kepada 101 pengguna aktif. Proses pembersihan data, encoding label, normalisasi Min-Max, dan pemilihan fitur adalah semua bagian dari proses pengolahan data. Dalam proses pemodelan, algoritma support vector machine (SVM) dengan kernel Radial Basis Function (RBF) digunakan sebagai metode klasifikasi. Pengembangan model analisis risiko keamanan yang bergantung pada persepsi dan perilaku pengguna daripada data log sistem atau transaksi aktual memberikan kontribusi unik penelitian ini. Metode ini menawarkan perspektif alternatif dalam identifikasi potensi risiko keamanan dengan menggunakan indikator non-teknis sebagai variabel prediktif pada platform e-commerce.

**Kata Kunci:** Analisis Perilaku, Keamanan Sistem Informasi, Machine Learning, Shopee, Support Vector Machine.

## ABSTRACT

The expansion of e-commerce in Indonesia has made information system security a crucial concern, especially on sites like Shopee that see a lot of user activity and transaction volumes. Potential security hazards, such as account misuse, unauthorized access, and suspicious activity, are increased by the volume of online transactions. Therefore, in order to comprehend the elements linked to security threats based on user characteristics and behavioral patterns, an analytical approach is necessary. The purpose of this study is to apply machine learning to examine security risk tendencies among Shopee users. A standardized questionnaire addressing demographic factors, usage frequency, security awareness levels, and experiences with questionable activity was used to gather data from 101 active users. Data cleaning, label encoding, Min–Max normalization, and feature selection were among the steps in the data processing procedure. The classification model used was the Support Vector Machine (SVM) technique with a Radial Basis Function (RBF) kernel. The creation of a security risk analysis model based on user perceptions and behavioral aspects rather than system log or transactional data is what makes this study unique. By using non-technical indications as predictive factors in e-commerce platforms, this method provides an alternate viewpoint for spotting possible security threats.

**Keywords:** Behavioral Analysis, Information System Security, Machine Learning, Shopee, Support Vector Machine

## I. PENDAHULUAN

Semakin Perkembangan teknologi digital telah mendorong peningkatan signifikan penggunaan platform e-commerce di Indonesia. E-commerce menjadi salah satu sektor dengan pertumbuhan tercepat dalam ekonomi digital [1] Shopee sebagai salah satu platform e-commerce terbesar di Asia Tenggara menunjukkan tingkat adopsi dan intensitas penggunaan yang

tinggi di Indonesia, yang tercermin dari dominasi Shopee dalam berbagai penelitian terkait perilaku konsumen dan niat pembelian Berdasarkan laporan [2], jumlah pengguna internet di Indonesia telah mencapai lebih dari 185 juta pengguna, dengan penetrasi internet sekitar 66% dari total populasi. Pertumbuhan ini berdampak langsung pada peningkatan aktivitas perdagangan elektronik (e-commerce) yang menjadi salah satu sektor utama dalam ekonomi digital Indonesia. [3] Tingginya aktivitas transaksi dan interaksi pengguna pada

platform e-commerce menyebabkan meningkatnya eksposur terhadap risiko keamanan dan privasi data. [4] menegaskan bahwa isu privasi data merupakan tantangan utama dalam ekosistem e-commerce modern, terutama terkait perlindungan informasi pribadi pengguna. Selain itu, [5] mengklasifikasikan ancaman siber pada *e-commerce* sebagai bentuk *Advanced Persistent Threat* (APT) yang terus berkembang dan beradaptasi. Dalam konteks Shopee, berbagai penelitian menunjukkan bahwa faktor keamanan, kualitas informasi, serta kepercayaan pengguna berpengaruh signifikan terhadap niat dan perilaku penggunaan aplikasi [3]. Dinamika kepercayaan ini juga dipengaruhi oleh mekanisme interaksi keamanan informasi berbasis perilaku [6] serta dinamika trust dalam adopsi sistem informasi [7]. Dengan jumlah pengguna aktif yang besar dan intensitas transaksi yang tinggi, potensi risiko seperti phishing, social engineering, percobaan login ilegal, penyalahgunaan akun, serta transaksi mencurigakan menjadi tantangan serius bagi keamanan sistem informasi. Oleh karena itu, diperlukan pendekatan analitis berbasis data untuk mengidentifikasi pola risiko secara lebih sistematis. Algoritma klasifikasi *Support Vector Machine* bekerja dengan menemukan hyperplane terbaik untuk memisahkan dua kelas data. SVM sangat baik untuk menangani data berdimensi besar dan kasus non-linear melalui penggunaan kernel trick. Kernel RBF (*Radial Basis Function*) adalah salah satu kernel paling populer karena dapat memetakan data ke ruang dimensi besar, yang membuat pemisahan pola non-linear lebih mudah. SVM sangat baik untuk penelitian berbasis kuesioner karena stabil pada dataset yang tidak terlalu besar. Selain itu, beberapa penelitian menunjukkan bahwa kinerja SVM sangat bergantung pada pemilihan parameter seperti nilai C dan gamma [8]

## II. STUDI PUSTAKA

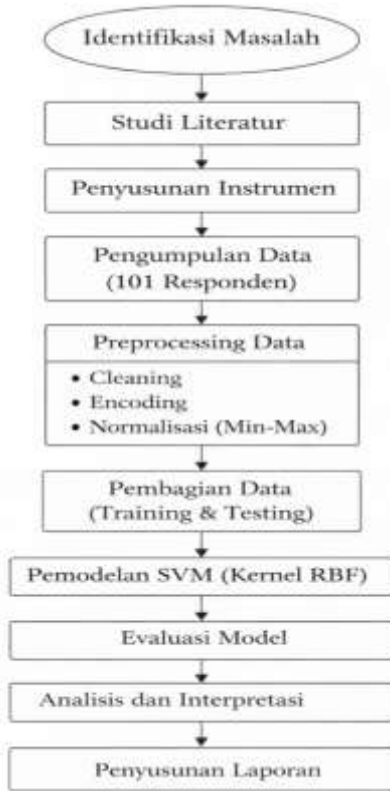
Berdasarkan masalah yang telah diuraikan di atas, serta merujuk pada penelitian sebelumnya terkait efektivitas algoritma klasifikasi, salah satu faktor yang berpengaruh dalam akurasi pengolahan data adalah penerapan teknologi *Support Vector Machine* (SVM). Algoritma ini bekerja dengan menemukan hyperplane terbaik untuk memisahkan kelas data. Penggunaan teknologi SVM sangat baik untuk menangani data berdimensi besar dan kasus non-linear melalui penggunaan kernel trick [8]. Penerapan Radial Basis Function (RBF) sebagai kernel utama diharapkan memudahkan pemisahan

pola non-linear karena kemampuannya memetakan data ke ruang dimensi yang lebih tinggi. Penggunaan SVM juga sangat stabil untuk penelitian berbasis kuesioner dengan dataset yang tidak terlalu besar, meskipun kinerjanya sangat bergantung pada pemilihan parameter  $C$  dan  $\gamma$ . Pemodelan klasifikasi risiko keamanan pada pengguna Shopee dapat menjadi solusi terbaik dari permasalahan ancaman siber di platform e-commerce. Penggunaan teknologi terbaru seperti SVM dapat membantu dalam mengidentifikasi tingkat potensi risiko secara akurat. Dengan adanya klasifikasi ini, diharapkan dapat memberikan gambaran bagi pengguna dalam mengenali aktivitas mencurigakan. Selanjutnya, model yang dibuat pada penelitian ini dapat dikembangkan lebih lanjut untuk meningkatkan sistem proteksi data pengguna. Pada penelitian sebelumnya, pernah dilakukan analisis sentimen terhadap pengguna Shopee menggunakan SVM, namun fokusnya terbatas pada opini publik dan tidak membahas risiko keamanan [9]. Tahun 2020 pernah dilakukan penelitian terkait perbandingan algoritma SVM dan Naive Bayes, tetapi studi tersebut tidak mengkaji pengalaman ancaman yang dialami pengguna secara langsung [10]. Penelitian lainnya juga mengevaluasi aspek keamanan dan minat beli, namun tidak melakukan klasifikasi terhadap tingkat risiko [10]. Selain itu, kajian mengenai trust dan keamanan sistem secara umum telah dilakukan, tetapi belum spesifik pada populasi pengguna Shopee di Indonesia [6]. Oleh karena itu, penelitian ini bertujuan mengisi celah tersebut dengan memfokuskan pada klasifikasi risiko keamanan yang lebih spesifik.

## III. METODE PENELITIAN

Pada penelitian ini penulis melakukan beberapa tahapan sebagaimana mestinya mulai dari penyusunan kemudian identifikasi masalah yang sudah Dalam melaksanakan penelitian ini, penulis menempuh serangkaian tahapan sistematis guna memastikan proses analisis data dan klasifikasi berjalan secara terstruktur serta akurat. Prosedur penelitian diawali dengan tahap penyusunan rencana penelitian yang mencakup penentuan batasan masalah dan tujuan yang ingin dicapai. Selanjutnya, dilakukan identifikasi masalah secara mendalam terhadap fenomena risiko keamanan pada platform e-commerce, khususnya bagi pengguna Shopee di Indonesia. Tahapan ini sangat krusial untuk memetakan variabel-variabel apa saja yang

akan digunakan dalam pemodelan *Support Vector Machine (SVM)* seperti gambar berikut ini



Gambar 1. Alur Penelitian

Instrumen penelitian ini menggunakan kuesioner yang terdiri atas tiga bagian utama, di mana bagian pertama bertujuan untuk mengumpulkan data demografi responden yang mencakup jenis kelamin, usia, dan lama menggunakan platform Shopee. Bagian kedua berfungsi untuk mengukur variabel persepsi dan perilaku pengguna dengan menggunakan instrumen Skala Likert 5 poin, yang dimulai dari skor 1 (Sangat Tidak Setuju) hingga skor 5 (Sangat Setuju). Penggunaan Skala Likert ini dipilih karena dinilai mampu mengukur sikap, persepsi, dan tingkat persetujuan responden secara kuantitatif terhadap empat variabel utama, yaitu persepsi keamanan, kewaspadaan penggunaan, pengalaman ancaman, serta perilaku penggunaan aplikasi. Terakhir, bagian ketiga merupakan variabel label atau target yang bersifat dikotomis (binary classification). Variabel target ini dikodekan dengan angka 1 untuk responden yang pernah mengalami aktivitas mencurigakan, dan angka 0 untuk responden yang tidak pernah mengalaminya, di mana variabel ini nantinya digunakan sebagai kelas target dalam proses klasifikasi menggunakan algoritma Support Vector Machine (SVM).

4. Uji Validitas dan Reliabilitas Instrumen. Sebelum dilakukan proses pemodelan, instrumen penelitian diuji untuk memastikan kualitas data melalui uji validitas dan uji reliabilitas. Uji validitas dilakukan menggunakan korelasi Pearson Product Moment antara skor item dengan skor total (item-total correlation).

Kriteria pengujian:

Jika  $r_{hitung} > r_{tabel} (\alpha = 0,05)$ , maka item dinyatakan valid.

Jika  $r_{hitung} \leq r_{tabel}$ , maka item dinyatakan tidak valid dan dieliminasi.

Dengan jumlah responden sebanyak 101 orang, nilai  $r_{tabel}$  pada taraf signifikansi 5% adalah sekitar 0,195.

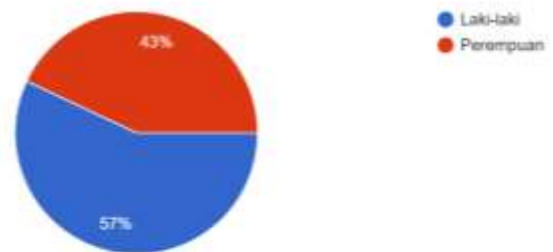
Hasil pengujian menunjukkan bahwa seluruh item pernyataan memiliki nilai  $r_{hitung}$  lebih besar dari  $r_{tabel}$ , sehingga seluruh item dinyatakan valid dan layak digunakan dalam penelitian.

#### 5. Uji Reliabilitas

Uji reliabilitas dilakukan menggunakan metode tim

### IV. HASIL DAN PEMBAHASAN

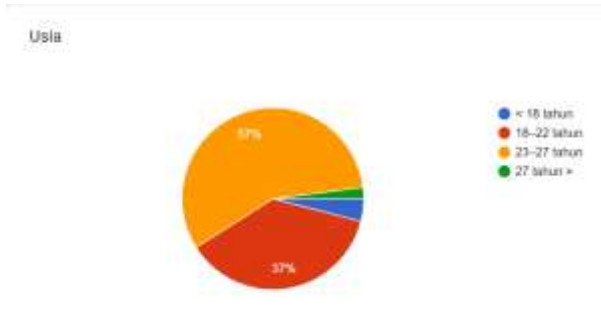
Penelitian ini melibatkan 101 responden pengguna aktif Shopee. Data demografi digunakan untuk memahami profil pengguna sekaligus melihat kecenderungan pola risiko keamanan berdasarkan karakteristik individu. Mayoritas responden berjenis kelamin perempuan, sedangkan laki-laki merupakan kelompok minoritas. Hal ini sejalan dengan studi e-commerce yang menunjukkan bahwa perempuan cenderung lebih aktif berbelanja secara daring dibanding laki-laki.



Gambar 3. Diagram Distribusi Jenis Kelamin Pengguna Shopee

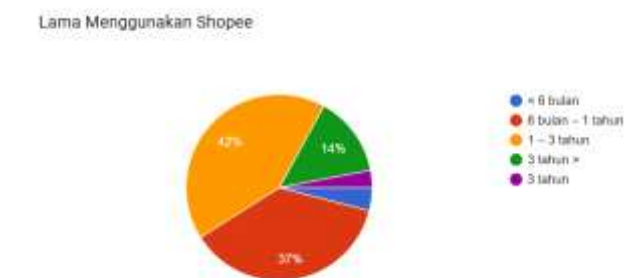
### 1) Distribusi Usia

Rentang usia responden didominasi kelompok **17–25 tahun**, yang merupakan pengguna aktif e-commerce di Indonesia. Kelompok usia ini memiliki tingkat penggunaan aplikasi tinggi, namun cenderung lebih rentan terhadap ancaman keamanan seperti phishing dan social engineering karena intensitas aktivitas digital yang besar.



Gambar 4. Diagram Rentang Usia Pengguna Shopee

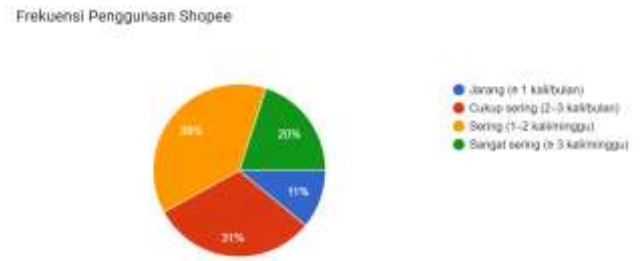
Sebagian besar responden sudah menggunakan Shopee selama **1–3 tahun**, menunjukkan tingkat pengalaman yang cukup dalam bertransaksi. Namun demikian, pengalaman panjang tidak selalu menjamin tingkat kewaspadaan pengguna terhadap ancaman digital.



Gambar 5. Diagram Lama Penggunaan Shopee oleh Responden

### 2) Frekuensi Penggunaan

Mayoritas responden menggunakan Shopee secara rutin setiap minggu, bahkan beberapa hampir setiap hari. Frekuensi penggunaan yang tinggi meningkatkan peluang terpapar ancaman seperti pesan phishing, promo palsu, atau akses login tidak sah.



Gambar 6. Diagram Frekuensi Penggunaan Shopee oleh Responden

### 3) Pengolahan Data dan Preprocessing

Sebelum pemodelan, data hasil kuesioner melalui beberapa tahap:

- Encoding label** (0 = tidak pernah mengalami aktivitas mencurigakan, 1 = pernah)
- Normalisasi Min-Max** untuk menyamakan skala seluruh fitur
- Pemilihan fitur** hanya untuk variabel numerik dan kategorikal
- Pemisahan target** berupa variabel pengalaman ancaman

Tahap preprocessing ini memastikan model SVM menerima data bersih, terstandarisasi, dan siap digunakan untuk proses klasifikasi.

		Predicted		
		0	1	$\Sigma$
Actual	0	2	24	26
	1	2	73	75
	$\Sigma$	4	97	101

Gambar 7. *Confusion Matrix* Hasil Evaluasi Model SVM

Hasil Pemodelan SVM Pemodelan dilakukan menggunakan aplikasi Orange Data Mining dengan algoritma SVM kernel RBF. Evaluasi

menggunakan 5-fold Cross Validation menghasilkan:

Metrik	Nilai
Accuracy	74.3%
AUC	0.757
Precision	0.688
Recall	0.743
F1-Score	0.665
MCC	0.113

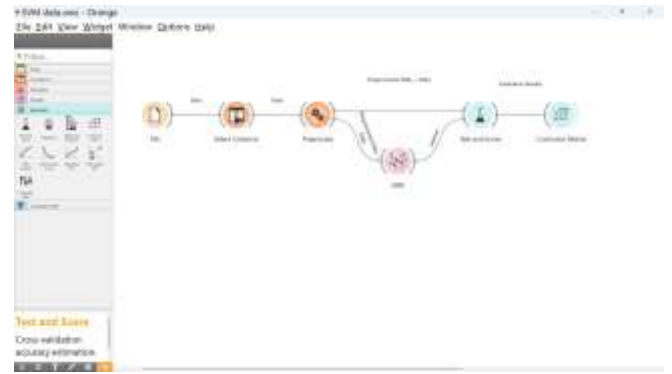
3) Interpretasi Hasil

Akurasi 74.3% menunjukkan model cukup baik dalam mengklasifikasikan pengguna yang pernah / tidak pernah mengalami aktivitas mencurigakan. AUC 0.757 berada pada kategori baik, menandakan bahwa SVM mampu membedakan dua kelas dengan cukup akurat. Precision 0.688 menunjukkan bahwa dari seluruh prediksi “pernah”, sekitar 68.8% benar. Recall 0.743 menunjukkan bahwa model berhasil menemukan sebagian besar responden yang benar-benar pernah mengalami aktivitas mencurigakan.

- A. **F1-score 0.665** menunjukkan keseimbangan yang moderat antara precision dan recall.
- B. **MCC 0.113** mengindikasikan hubungan prediksi-label yang lemah, kemungkinan dipengaruhi oleh ketidakseimbangan jumlah data antara kelas “pernah” dan “tidak pernah”.



Gambar 8. Hasil evaluasi model SVM menggunakan Test & Score di Orange Data Mining.



Gambar 9. Hasil Evaluasi Model SVM

B. Analisis Confusion Matrix

Hasil Confusion Matrix menunjukkan performa klasifikasi model sebagai berikut:

- **True Positive (TP): 32**
- Model berhasil mengidentifikasi 32 responden yang benar-benar pernah mengalami aktivitas mencurigakan.
- **True Negative (TN): 43**
- Model mampu mengklasifikasikan dengan benar 43 responden yang tidak pernah mengalami aktivitas mencurigakan.
- **False Positive (FP): 15**
- Terdapat 15 responden yang diprediksi pernah mengalami aktivitas mencurigakan, padahal sebenarnya tidak.
- **False Negative (FN): 11**
- Sebanyak 11 responden yang pernah mengalami aktivitas mencurigakan tidak terdeteksi oleh model.

Berdasarkan hasil tersebut, model menunjukkan kemampuan yang cukup baik dalam mengidentifikasi kedua kelas, meskipun masih terdapat kesalahan klasifikasi terutama pada kategori False Positive dan False Negative yang memengaruhi nilai precision dan recall. FN memerlukan perhatian karena kesalahan tipe ini berpotensi mengurangi efektivitas sistem keamanan jika model digunakan sebagai deteksi dini.

Hasil penelitian menunjukkan bahwa metode SVM efektif dalam menganalisis faktor-faktor yang mempengaruhi pengalaman pengguna terhadap aktivitas mencurigakan di platform Shopee. Meskipun akurasi model tidak mencapai tingkat yang sangat tinggi, nilai AUC di atas 0.75 menunjukkan kemampuan model dalam membedakan pola risiko. Meskipun model ini memberikan

wawasan yang berharga, terdapat beberapa keterbatasan teknis yang perlu diperhatikan. Dataset yang terbatas pada 101 responden menyebabkan pola klasifikasi belum mencapai potensi maksimalnya, ditambah dengan adanya ketidakseimbangan kelas—di mana mayoritas responden belum pernah mengalami ancaman—yang menurunkan stabilitas metrik Matthews Correlation Coefficient (MCC). Selain itu, penggunaan fitur berbasis kuesioner membawa tingkat subjektivitas yang lebih tinggi dibandingkan dengan data perilaku objektif. Namun demikian, model ini tetap berhasil menunjukkan pola kecenderungan yang konsisten: pengguna dengan frekuensi aktivitas tinggi yang tidak dibarengi dengan wawasan keamanan yang memadai memiliki risiko yang jauh lebih besar terhadap ancaman keamanan. Secara keseluruhan, SVM memberikan hasil yang cukup baik sebagai model awal dalam mendeteksi potensi risiko keamanan berdasarkan persepsi pengguna.

## V. KESIMPULAN

Tujuan penelitian ini adalah untuk mengevaluasi keamanan sistem informasi pengguna Shopee dengan menggunakan algoritma Support Vector Machine (SVM), yang berfungsi sebagai metode klasifikasi data perilaku dan atribut pengguna. Beberapa temuan penting dibuat berdasarkan hasil pengolahan data melalui proses preprocessing, pemilihan atribut, dan evaluasi model melalui cross-validation.

Pertama, analisis demografis menunjukkan bahwa sebagian besar pengguna Shopee berasal dari kelompok usia produktif (23–27 tahun) dan menggunakan aplikasi dengan frekuensi yang cukup tinggi. Ini menunjukkan bahwa responden memiliki pengalaman yang cukup untuk menilai dan menunjukkan pola perilaku terkait keamanan penggunaan aplikasi.

Selanjutnya, hasil pengujian model SVM menunjukkan nilai kinerja yang cukup baik: AUC sebesar 0,757, akurasi sebesar 0,743, ketepatan sebesar 0,688, recall sebesar 0,743, dan F1-Score sebesar 0,665. Nilai-nilai ini menunjukkan bahwa model dapat dengan sukses mengkategorikan perilaku pengguna yang berkaitan dengan indikator keamanan. Meskipun tidak memiliki akurasi yang tinggi, kinerja model menunjukkan bahwa SVM dapat mendeteksi pola-pola yang berhubungan dengan risiko keamanan dengan baik.

Ketiga, hasil Confusion Matrix menunjukkan bahwa model SVM dapat mengidentifikasi kelas mayoritas dengan baik, meskipun beberapa kelas

minor memiliki kesalahan klasifikasi. Ini menunjukkan bahwa ada perluasan kualitas data seperti mengimbangi jumlah sampel, menambah variabel yang lebih representatif, atau menggunakan metode tuning parameter untuk meningkatkan kinerja model.

Secara keseluruhan, penelitian ini menunjukkan bahwa Support Vector Machine dapat digunakan sebagai alat analisis awal untuk mengidentifikasi pola keamanan pengguna Shopee, terutama yang berkaitan dengan aktivitas penggunaan aplikasi. Untuk mendapatkan hasil yang lebih komprehensif, penelitian selanjutnya dapat mempertimbangkan teknik pengembangan fitur yang lebih mendalam, penambahan jumlah responden, dan perbandingan dengan algoritma lain.

## DAFTAR PUSTAKA

- [1] A. Kumar, "E - commerce in India," vol. 4, no. 2, pp. 103–105, 2021.
- [2] W. A. Social and Meltwater, "Digital 2024 Indonesia, THE ESSENTIAL GUIDE TO THE LATEST CONNECTED BEHAVIOURS," p. 136, 2024, [Online]. Available: <https://n9.cl/5udw2>
- [3] G. P. Mandiri and I. Susila, "Analysis of the Effect of Perceived Security, Information Quality, and Consumer Trust on Online Purchase Intention in the Shopee Application," *JEM17 J. Ekon. Manaj.*, vol. 8, no. 2, pp. 241–260, 2023, doi: 10.30996/jem17.v8i2.9765.
- [4] D. Arora, "Data Privacy Issues With E-Commerce," *Int. J. Soc. Sci. Econ. Res.*, vol. 08, no. 05, pp. 1167–1174, 2023, doi: 10.46609/ijsser.2023.v08i05.020.
- [5] H. Desamsetti, "Crime and Cybersecurity as Advanced Persistent Threat: A Constant E-Commerce Challenges," *Am. J. Trade Policy*, vol. 8, no. 3, pp. 239–246, 2021, doi: 10.18034/ajtp.v8i3.666.
- [6] J. Gao, H. Yu, H. Wang, H. Bai, C. Huo, and G. Zhang, "Research on information security interaction based on behavior trust measurement mechanism," *ITM Web Conf.*, vol. 47, p. 03015, 2022, doi: 10.1051/itmconf/20224703015.
- [7] L. S. Müller, C. Nohe, S. Reiners, J. Becker, and G. Hertel, "Adopting information systems at work: a longitudinal examination of trust dynamics, antecedents, and outcomes," *Behav. Inf. Technol.*, vol. 43, no. 6, pp. 1096–1128, 2024, doi: 10.1080/0144929X.2023.2196598.
- [8] H. H. Huang, Z. Wang, and W. Chung, "Efficient parameter selection for support vector machines," *Enterp. Inf. Syst.*, vol. 13, no. 6, pp. 916–932, 2019, doi: 10.1080/17517575.2019.1592233.
- [9] Idris et al, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)boards," *J. Inf. Sci.*, vol. 36, no. 6, pp. 823–848, 2023.
- [10] P. S. Hutapea and W. Maharani, "Sentiment Analysis

on Twitter Social Media towards Shopee E-Commerce through Support Vector Machine (SVM) Method,” *JINAV J. Inf. Vis.*, vol. 4, no. 1, pp. 7–17, 2023, doi: 10.35877/454ri.jinav1504.