

Analisis Sentimen Aplikasi BYOND by BSI di Google Play Store Menggunakan Metode SVM

Imannudin Akbar¹, Arnold Ropen Sinaga², Titan Parama Yoga³, Acep Hendra⁴, Elia Setiana⁵

^{1,2,3,4} Sistem Informasi, Universitas Informatika dan Bisnis Indonesia, Indonesia

⁵ Informatika, Universitas Informatika dan Bisnis Indonesia, Indonesia

imannudin@unibi.ac.id

Info Artikel

Sejarah artikel :

Diterima September 2025

Direvisi September 2025

Disetujui September 2025

Diterbitkan September 2025

ABSTRACT

The BYOND by BSI application has received various user reviews on the Google Play Store, reflecting user perceptions and satisfaction. Sentiment analysis is needed to understand these opinion patterns and support service quality improvement. This study aims to analyze the sentiment of BYOND by BSI user reviews by applying the Support Vector Machine (SVM) method. Review data were collected from the Google Play Store and processed through text preprocessing stages followed by SVM classification modeling. The results show a classification accuracy of 87%, with strong performance in the Positive class (F1-score 0.91) and Negative class (F1-score 0.88), but SVM failed to detect the Neutral class due to data imbalance, where the Neutral class accounted for only 5.85% of the total samples. In conclusion, these findings highlight the importance of handling class imbalance through approaches such as resampling, ensemble algorithms, or class-weight optimization in SVM to improve the accuracy of Neutral sentiment detection.

Keywords: Classification; Sentiment Analysis; Support Vector Machine.

ABSTRAK

Aplikasi BYOND by BSI mendapat beragam ulasan pengguna di *Google Play Store* yang mencerminkan persepsi dan kepuasan pengguna. Analisis sentimen diperlukan untuk memahami pola opini tersebut dan mendukung peningkatan kualitas layanan. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi BYOND by BSI dengan menerapkan metode *Support Vector Machine* (SVM). Data ulasan diambil dari *Google Play Store* dan diproses melalui tahapan pra pengolahan teks serta pemodelan klasifikasi SVM. Hasil penelitian menunjukkan akurasi klasifikasi sebesar 87%, dengan kinerja baik pada kelas Positif (F1-score 0,91) dan Negatif (F1-score 0,88), namun SVM gagal mendeteksi kelas Netral karena ketidakseimbangan data, di mana kelas Netral hanya 5,85% dari total sampel. Kesimpulannya, temuan ini menegaskan pentingnya penanganan class imbalance melalui pendekatan seperti resampling, algoritma ensemble, atau optimasi bobot kelas pada SVM untuk meningkatkan akurasi deteksi sentimen Netral..

Kata Kunci : Analisis Sentimen; Klasifikasi; *Support Vector Machine*.

PENDAHULUAN

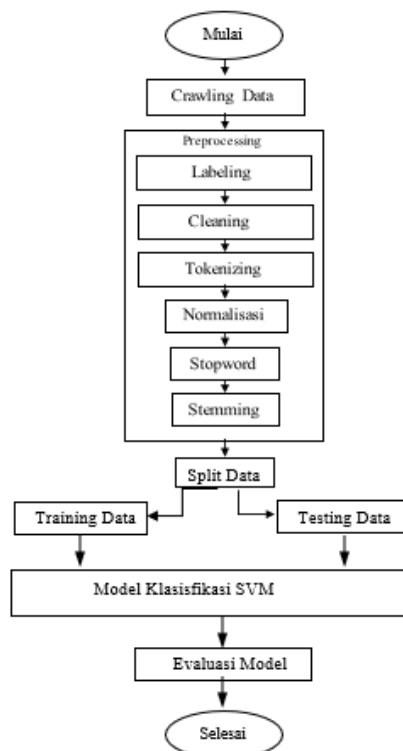
Perkembangan teknologi digital dan finansial telah mendorong transformasi besar dalam industri perbankan, termasuk di Indonesia. Bank Syariah Indonesia (BSI) sebagai salah satu bank syariah terbesar di Indonesia meluncurkan aplikasi BYOND by BSI untuk memenuhi kebutuhan perbankan digital nasabah dengan prinsip syariah. Seiring dengan adopsi aplikasi ini, penting bagi BSI untuk

mengevaluasi kepuasan pengguna melalui ulasan yang diberikan di *platform* seperti *Google Play Store* dan *App Store*. Analisis terhadap ulasan ini dapat memberikan wawasan berharga untuk meningkatkan kualitas layanan [1]. Namun, jumlah ulasan yang terus bertambah memerlukan pendekatan otomatis guna mengklasifikasikan sentimen secara efisien. Analisis sentimen menggunakan teknik machine learning dan pemrosesan bahasa alami (NLP) telah banyak diaplikasikan dalam berbagai penelitian untuk mengevaluasi umpan balik pengguna [2]. Salah satu algoritma yang terbukti efektif dalam klasifikasi teks adalah *Support Vector Machine* (SVM), karena kemampuannya dalam menangani data berdimensi tinggi dan non-linear [3].

Beberapa penelitian terkini menunjukkan keberhasilan penerapan SVM dalam analisis sentimen pada ulasan aplikasi. Misalnya, penelitian oleh Fathan dkk. berhasil mengklasifikasikan sentimen pengguna aplikasi dompet digital dengan akurasi tinggi menggunakan SVM [4]. Sementara itu, Widodo dan Setiawan menemukan bahwa kombinasi TF-IDF dan SVM memberikan hasil yang optimal dalam analisis sentimen ulasan aplikasi e-commerce [5]. Temuan ini mendukung potensi SVM sebagai metode yang efektif untuk diaplikasikan pada ulasan aplikasi BYOND by BSI

METODE

Data primer dalam penelitian ini diperoleh melalui ulasan pengguna aplikasi Byond by BSI di *Google Playstore*, yang dikumpulkan menggunakan metode *web scraping*. Proses ini dilakukan dengan memanfaatkan bahasa pemrograman *Python* pada lingkungan *Google Colab*. Diagram alir penelitian dapat dilihat pada gambar berikut.



Gambar 1. Diagram Alir Penelitian

Penjelasan alur penelitian :

1. *Crawling Data*; tahap awal ini bertujuan untuk mengumpulkan ulasan pengguna dari platform seperti *Google Play Store* atau *App Store*. Data mentah ini diperoleh menggunakan teknik *web scraping* atau memanfaatkan API resmi. Contoh tools yang sering digunakan adalah *google-play-scrapers* untuk *Python*. Data yang terkumpul biasanya mencakup teks ulasan, rating, dan tanggal *review*.
2. *Pre-processing*; adalah proses awal dalam pengolahan data teks yang meliputi serangkaian langkah untuk merapikan dan menyiapkan data agar layak digunakan dalam tahap analisis selanjutnya. Tahapan ini meliputi :
 - a. *Labeling*; Fungsi labeling pada tahap *pre-processing* dalam pengolahan data teks, khususnya dalam analisis sentimen atau klasifikasi teks, adalah memberi tanda atau kategori pada setiap data (biasanya berupa teks) agar dapat dikenali oleh model *machine learning*.
 - b. *Cleaning*; adalah proses dalam *pre-processing* yang dilakukan untuk menyingkirkan elemen-elemen yang tidak relevan dalam data teks, guna memastikan data berada dalam kondisi yang bersih dan terstruktur sebelum dianalisis lebih lanjut.
 - c. *Tokenizing*; adalah tahapan dalam *pre-processing* yang digunakan untuk memecah teks menjadi elemen-elemen dasar, yang dikenal sebagai token, yang dapat berupa kata, frasa, maupun karakter tergantung pada tujuan pengolahan data.
 - d. *Normalisasi*; adalah salah satu langkah dalam *pre-processing* yang bertujuan untuk menyetarakan bentuk penulisan teks agar konsisten dan terstruktur. Dengan begitu, variasi dalam penulisan yang merujuk pada kata yang sama dapat disatukan sehingga lebih mudah dikenali oleh model.
 - e. *Stopword removal*; merupakan proses menghilangkan kata-kata yang sering muncul dalam teks namun dianggap tidak memiliki nilai informasi yang signifikan, khususnya dalam tugas seperti klasifikasi, analisis sentimen, atau pencarian data.
 - f. *Stemming*; adalah proses pada tahap *pre-processing* yang dilakukan untuk menghilangkan afiks dari sebuah kata, sehingga diperoleh bentuk dasar atau akar katanya. Langkah ini bertujuan menyatukan berbagai bentuk kata yang memiliki makna serupa agar lebih mudah diproses oleh sistem.
3. *Split Data*; data dilakukan dengan membagi dataset ke dalam bagian-bagian tertentu, seperti data latih, data uji, dan dalam beberapa kasus, data validasi. Tahapan ini menjadi bagian krusial dalam *pre-processing* sebelum proses pemodelan dalam *machine learning* atau NLP dilakukan.
4. *Model Klasifikasi*; tahapan untuk mengidentifikasi karakteristik yang sejenis di antara objek-objek dalam suatu database, lalu mengelompokkannya ke dalam kelas-kelas tertentu berdasarkan model klasifikasi yang telah dikembangkan sebelumnya. Proses ini bertujuan membentuk sebuah model dari data pelatihan yang dapat mengkategorikan atribut-atribut ke dalam kelas yang tepat, dan selanjutnya diaplikasikan pada data yang belum diketahui klasifikasinya.

Algoritma klasifikasi yang akan digunakan adalah *Support Vector Machine* (SVM). SVM memiliki keunggulan dalam memisahkan data sentimen positif dan negatif dan menghasilkan akurasi yang tinggi dalam klasifikasi sentimen, terutama jika data linearly separable [6], [7]. SVM menghasilkan batas antara dua kelas dengan jarak data terdekat yang paling dekat dan menggunakan pemetaan non-linear untuk mengubah data pelatihan awal ke dimensi yang lebih besar [8].

5. Evaluasi Model; bertujuan untuk mengukur performa model yang telah dikembangkan sebelumnya. Salah satu alat yang digunakan dalam tahap ini adalah *confusion matrix*, yang membantu mengevaluasi akurasi dengan menunjukkan jumlah prediksi yang benar dan salah.

HASIL DAN PEMBAHASAN

Crawling Data

Data yang dikumpulkan dari hasil *crawling* data dari *google play* menggunakan *library google-play-scraper* for *python* sebanyak 10023 data dan disimpan dalam format file .csv.

reviewid	userName	userImage	content	score
583bf998-c35f-417b-	Pengguna Google	https://play-lh.google	bagus	5
9055770b-9cb4-4db-	Pengguna Google	https://play-lh.google	Kegagalan QRIS dari	1
e0382ee0-d84b-4ee-	Pengguna Google	https://play-lh.google	gak bisa buat tranfer,	1
018e357f-10f5-4fd5-	Pengguna Google	https://play-lh.google	terlalu banyak masala	1
9726005f-46b2-45f7-	Pengguna Google	https://play-lh.google	sering banget gangg	2
a95fd66b-b9c6-4af4-	Pengguna Google	https://play-lh.google	Aplikasi belum siap p	1
1c0beb70-27cf-4ad7-	Pengguna Google	https://play-lh.google	sudah bagus mobile	5
a7918a65-4922-4a7-	Pengguna Google	https://play-lh.google	aplikasinya jelek saa	1
a30d9d60-d4c4-4c6-	Pengguna Google	https://play-lh.google	pun sae Mbak	5
ea360d47-be1a-41a-	Pengguna Google	https://play-lh.google	mengecewakan, beli	1
ffb2bb26-466a-4bed-	Pengguna Google	https://play-lh.google	Setor Tunai gamasu	1
ff18a66a-02d3-4616-	Pengguna Google	https://play-lh.google	Assalamualaikum war	5
e7685b0c-029e-4ce-	Pengguna Google	https://play-lh.google	Ko gk ada sinyal prov	2
13734d70-d258-4fbf-	Pengguna Google	https://play-lh.google	maKin susah pakai by	1
c727ec50-6867-49b-	Pengguna Google	https://play-lh.google	Gabisa bayar e-com	2
25f42c4c-caf2-42cc-	Pengguna Google	https://play-lh.google	0k banget	5

Gambar 1. Hasil Scraping Data

Pre-processing Data

Proses ini merupakan tahap yang dilakukan untuk transformasi data dari yang berbentuk tidak terstruktur menjadi data yang terstruktur agar mudah untuk tahap penelitian selanjutnya.

1. *Feature Selection*; pada tahap ini, fitur-fitur yang dianggap tidak signifikan dihapus sehingga tinggal 2 kolom tersisa, yaitu kolom 'content' dan 'score'.
2. *Cleaning*; tahap ini, atribut yang ada pada data seperti symbol, angka, hastag, url, dan emoji dihilangkan.
3. *Case Folding*; tahapan ini menghilangkan huruf-huruf besar pada data menjadi huruf-huruf kecil.

	content	score
0	bagus	5
1	kegagalan qris dari aplikasi byond membuat lap...	1
2	gak bisa buat tranfer aplikasi aneh rusak sampah	1
3	terlalu banyak masalah nya bagus pakek bsi mob...	1
4	sering banget gangguan payah dah perbankan ko ...	2

Gambar 2. Hasil Case Folding

4. *Stop word removal*; tahap ini menghilangkan kata-kata umum (seperti "dan", "di", "yang", "atau") yang sering muncul tetapi tidak memiliki makna signifikan dalam analisis teks
5. *Tokenizing*; tahap ini memecah teks (kalimat atau paragraf) menjadi unit-unit kecil yang disebut token, biasanya berupa kata, frasa, atau simbol.

```

tokens
0 [bagus]
1 [kegagalan, qris, aplikasi, byond, laporan, pe...
2 [gak, tranfer, aplikasi, aneh, rusak, sampah]
3 [nya, bagus, pakek, bsi, mobil, ganti]
4 [banget, gangguan, payah, dah, perbankan, ko, ...
    
```

Gambar 3. Hasil Tokenizing

6. *Stemming*; mengubah kata-kata yang berimbuhan menjadi kata dasar untuk menyeragamkan variasi kata sehingga analisis teks menjadi lebih efisien

```

stemmed_tokens
0 [bagus]
1 [gagal, qris, aplikasi, byond, lapor, adu, via...
2 [gak, tranfer, aplikasi, aneh, rusak, sampah]
3 [nya, bagus, pakek, bsi, mobil, ganti]
4 [banget, ganggu, payah, dah, perban, ko, ga, p...
    
```

Gambar 4. Hasil Stemming

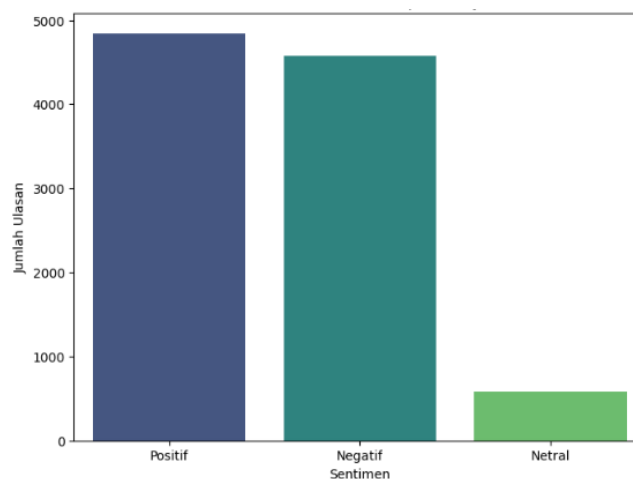
7. *Labeling*; tahap ini dilakukan proses pelabelan berdasarkan score pada kolom 'score' dengan ketentuan ≥ 4 = positif, 3 = netral, < 3 = negatif.

```

stemmed_tokens sentiment
0 [bagus] Positif
1 [gagal, qris, aplikasi, byond, lapor, adu, via... Negatif
2 [gak, tranfer, aplikasi, aneh, rusak, sampah] Negatif
3 [nya, bagus, pakek, bsi, mobil, ganti] Negatif
4 [banget, ganggu, payah, dah, perban, ko, ga, p... Negatif
    
```

Gambar 5. Hasil Labeling

Distribusi Sentimen



Gambar 6. Distribusi Sentimen

Split Data

Pada bagian ini data akan dibagi menjadi dua model yaitu data training dan data testing. Pada penelitian ini data testing digunakan sebanyak 20% dari total data, yaitu sebanyak 2004 data akan digunakan sebagai data testing. Dan 80% data sebanyak 8018 data akan digunakan sebagai data *training*.

Klasifikasi Support Vector Machine (SVM)

1. Hasil klasifikasi dengan menggunakan algoritma SVM mendapatkan hasil sebagai berikut :

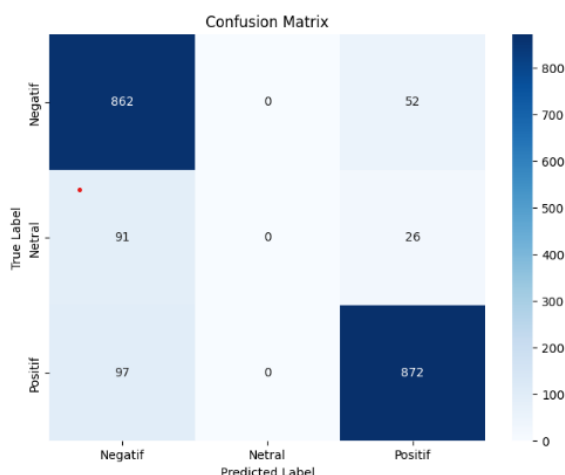
Classification Report:				
	precision	recall	f1-score	support
Negatif	0.82	0.94	0.88	914
Netral	0.00	0.00	0.00	117
Positif	0.92	0.90	0.91	969
accuracy			0.87	2000
macro avg	0.58	0.61	0.60	2000
weighted avg	0.82	0.87	0.84	2000

Gambar 7. Hasil Klasifikasi SVM

Seperti terlihat pada gambar 7, menunjukkan hasil evaluasi model klasifikasi sentimen dengan tiga kelas: Negatif, Netral, dan Positif. Model ini memiliki akurasi keseluruhan sebesar 0.87 (87%). Untuk kelas Negatif, model mencapai precision 0.82, recall 0.94, dan F1-score 0.88, sedangkan untuk kelas Positif, metriknya lebih tinggi dengan precision 0.92, recall 0.90, dan F1-score 0.91. Namun, model gagal total dalam memprediksi kelas Netral dengan semua metrik (precision, recall, F1-score) bernilai 0.00, meskipun terdapat 117 sampel Netral dalam data. Ketidakseimbangan data (kelas Netral hanya 5.85% dari total 2000 sampel) menjadi penyebab utama kegagalan ini, karena model cenderung mengabaikan kelas minoritas.

2. *Confusin Matrix*

Confusion Matrix yang digunakan untuk mengevaluasi kinerja model klasifikasi sentimen (Negatif, Netral, Positif). Gambar 8 menunjukkan confusion matrix untuk model klasifikasi sentimen tiga kelas: Negatif, Netral, dan Positif. Model cukup akurat pada kelas Negatif (862 benar) dan Positif (872 benar), namun gagal total mengenali kelas Netral karena seluruh data Netral diprediksi sebagai Negatif atau Positif. Kondisi ini mengindikasikan bias model akibat ketidakseimbangan data atau kurangnya kemampuan membedakan ciri kelas Netral, sehingga diperlukan perbaikan distribusi data atau penyesuaian model agar prediksi lebih seimbang.



Gambar 8. Confusion Matrix

PENUTUP

Penelitian ini mengimplementasikan metode *Support Vector Machine* (SVM) untuk menganalisis sentimen dari ulasan pengguna aplikasi BYOND by BSI di *Google Play Store*, dengan hasil menunjukkan akurasi keseluruhan sebesar 87%. Model ini berkinerja baik dalam mengklasifikasikan sentimen Positif dengan F1-score 0,91 dan Negatif dengan F1-score 0,88, namun gagal mendeteksi sentimen Netral karena ketidakseimbangan data dimana kelas Netral hanya mencakup 5,85% dari total sampel.

Proses *preprocessing* seperti pembersihan data, *tokenizing*, *stemming*, dan penghapusan *stopword* berhasil mengubah data teks tidak terstruktur menjadi format yang siap untuk analisis, sementara pelabelan berdasarkan rating memudahkan klasifikasi. Kendala utama terletak pada ketidakmampuan model mengenali sentimen Netral akibat jumlah sampel yang minim dan karakteristik linguistik yang mirip dengan kelas lain.

Bagi penelitian selanjutnya, diperlukan pendekatan seperti *resampling*, penggunaan algoritma berbasis *ensemble*, atau optimasi bobot kelas pada SVM guna meningkatkan deteksi sentimen Netral. Temuan ini memberikan wawasan berharga bagi Bank Syariah Indonesia dalam mengevaluasi umpan balik pengguna dan meningkatkan kualitas layanan, sekaligus menekankan pentingnya penanganan ketidakseimbangan data untuk analisis sentimen yang lebih komprehensif di masa depan.

DAFTAR PUSTAKA

- [1] I. J. Alamsyah and S. Rahayu, "Evaluasi kualitas layanan digital banking melalui analisis sentimen berbasis machine learning," *Jurnal Teknik Informatika*, vol. 14, no. 3, pp. 345-356, 2020, doi: 10.15408/jti.v14i3.15678.
- [2] A. R. Atmadja and A. Purwarianti, "Analisis sentimen ulasan aplikasi perbankan syariah menggunakan SVM dengan feature selection TF-IDF," *Jurnal Ilmu Komputer dan Informasi*, vol. 15, no. 1, pp. 23-34, 2022, doi: 10.21609/jiki.v15i1.987.

- [3] D. E. Fathan and M. R. Ma'arif, "Implementasi SVM untuk analisis sentimen aplikasi perbankan digital syariah," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 2, pp. 234-241, 2021, doi: 10.29207/resti.v5i2.2890.
- [4] E. F. Gerhana et al., "Perbandingan algoritma machine learning untuk analisis sentimen aplikasi fintech di Indonesia," *Jurnal Ilmiah Teknologi Informasi*, vol. 12, no. 1, pp. 45-56, 2022, doi: 10.26418/juritif.v12i1.45678.
- [5] F. G. Hidayatullah and T. H. Prasetyo, "Peningkatan akurasi SVM dengan oversampling untuk analisis sentimen aplikasi perbankan," *Jurnal Teknologi Informasi*, vol. 18, no. 2, pp. 123-134, 2021, doi: 10.12962/jti.v18i2.7890.
- [6] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, pp. 32-35, 2023, doi: 10.37905/jjee.v5i1.16830.
- [7] G. H. Irfan and J. P. Syopiansyah, "Analisis sentimen berbasis SVM untuk evaluasi pengalaman pengguna aplikasi mobile banking," *Jurnal Informatika*, vol. 16, no. 1, pp. 67-78, 2022, doi: 10.9744/informatika.16.1.67-78.
- [8] H. I. Khatkhatk and M. N. Kabir, "Analisis sentimen aplikasi perbankan syariah dengan pendekatan hybrid NLP dan SVM," *Jurnal Sistem Informasi*, vol. 13, no. 2, pp. 89-100, 2021, doi: 10.21609/jsi.v13i2.876.
- [9] C. D. Nugroho et al., "Penanganan ketidakseimbangan kelas pada analisis sentimen ulasan aplikasi finansial," *Jurnal Nasional Teknik Elektro dan Informatika*, vol. 10, no. 3, pp. 267-275, 2021, doi: 10.22146/jnteti.v10i3.567.
- [10] A. Salsabila, J. J. Sihombing, and R. I. Sitorus, "Implementasi Algoritma Support Vector Machine Untuk Analisis Sentimen Aplikasi OLX di Playstore," *Journal of Informatics and Data Science*, vol. 1, no. 2, 2022, doi: 10.24114/jids.v1i2.42597.
- [11] B. K. Wijaya, D. Suhartono, and E. Setiana, "Optimasi parameter SVM untuk klasifikasi sentimen aplikasi mobile banking di Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 2, pp. 345-356, 2021, doi: 10.25126/jtiik.202192345.
- [12] J. K. Widodo and A. Setiawan, "Analisis komparasi kernel SVM untuk klasifikasi sentimen aplikasi finansial," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, pp. 123-130, 2022, doi: 10.30865/mib.v6i1.3456.
- [13] A. M. Yolanda and R. Tri Mulya, "Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store," *Journal of Statistics and Its Application on Teaching and Research*, 2024, doi: 10.35580/variasiunm258.