

Sentiment Analysis of Fans Toward Brand Merchandise Releases Using Support Vector Machine (SVM)

Christian Imanuel Munaiseche¹, Nurchim², Bangun Prajadi Cipto Utomo³

^{1,2,3} Teknik Informatika, Ilmu Komputer, Universitas Duta Bangsa, Surakarta, 57154, Indonesia

Informasi Artikel

Diterima : 07 Juni 2025
Revisi : 27 Juni 2025
Publikasi : 30 September 2025

Kata Kunci:

Analisis Sentimen
Support Vector Machine
TF-IDF
Media Sosial
SMOTE

ABSTRAK

Analisis sentimen berbasis teks merupakan salah satu tantangan penting dalam memahami opini publik di media sosial, terutama dalam konteks komunitas penggemar idola. Ketidakterstruktur data teks dan ketidakseimbangan distribusi sentimen menjadi kendala utama dalam memperoleh hasil klasifikasi yang akurat. Penelitian ini bertujuan untuk menerapkan analisis sentimen terhadap respons publik di media sosial, khususnya terkait perilsan merchandise ulang tahun anggota JKT48 di platform X (Twitter), dengan memanfaatkan algoritma Support Vector Machine (SVM) yang efektif untuk klasifikasi teks berdimensi tinggi, serta teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk mengatasi ketidakseimbangan data. Sebanyak 1.062 komentar berhasil dikumpulkan menggunakan Tweet Harvest dan dilabeli ke dalam tiga kategori sentimen: positif, netral, dan negatif. Tahapan pre-processing mencakup case folding, cleansing, tokenisasi, dan stopword removal, dilanjutkan dengan ekstraksi fitur menggunakan metode TF-IDF. Hasil pengujian menunjukkan bahwa model SVM tanpa SMOTE memperoleh akurasi sebesar 84,62% dan F1-score sebesar 76,79%, sementara setelah penerapan SMOTE, akurasi meningkat menjadi 90,09% dan F1-score menjadi 90,15%. Pengujian tambahan menggunakan 5-fold cross-validation juga menunjukkan adanya peningkatan kinerja model secara keseluruhan. Penelitian ini berkontribusi terhadap pengembangan penerapan Natural Language Processing (NLP) dalam analisis sentimen berbasis teks di media sosial, khususnya untuk memahami persepsi publik pada komunitas penggemar idola di Indonesia.

ABSTRACT

Text-based sentiment analysis presents a significant challenge in understanding public opinion on social media, particularly within idol fan communities. The unstructured nature of text data and the imbalance in sentiment distribution are major obstacles to achieving accurate classification results. This study aims to apply sentiment analysis to public responses on social media, specifically related to the release of birthday merchandise for JKT48 members on the X (formerly Twitter) platform, by utilizing the Support Vector Machine (SVM) algorithm, which is known for its effectiveness in handling high-dimensional text classification tasks, and the Synthetic Minority Over-sampling Technique (SMOTE) to address data imbalance issues. A total of 1,062 comments were collected using Tweet Harvest and manually labeled into three sentiment categories: positive, neutral, and negative. The pre-processing stages included case folding, cleansing, tokenization, and stopword removal, followed by feature extraction using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The experimental results show that the SVM model without SMOTE achieved an accuracy of 84.62% and an F1-score of 76.79%, while after applying SMOTE, the accuracy increased to 90.09% and the F1-score to 90.15%. Additional testing using 5-fold cross-validation also demonstrated an overall improvement in model performance. This study contributes to the development of Natural Language Processing (NLP) applications in sentiment analysis for social media, particularly in understanding public perception within Indonesian idol fan communities.

***Penulis Koresponden**Email: 210103128@mhs.udb.ac.id

Cara sitasi IEEE:

C.I.Munaiseche, Nurchim, dan B.P.C. Utomo, "Sentiment Analysis of Fans Toward Brand Merchandise Releases Using Support Vector Machine (SVM)," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 3, pp. 1008-1013, September 2025, doi: 10.30811/jaise.v5i3.7264

1. PENDAHULUAN

Di era digital saat ini, industri hiburan tidak lagi hanya bergantung pada media konvensional untuk membangun citra, melainkan sangat dipengaruhi oleh opini publik yang tersebar luas di media sosial. Platform seperti X (sebelumnya Twitter) telah menjadi "panggung" utama bagi para penggemar untuk menyuarakan dukungan, kritik, bahkan kekecewaan mereka terhadap produk hiburan atau figur publik. Ketika kontroversi terjadi, opini yang berkembang di media sosial dapat menyebar dalam hitungan menit dan berdampak besar terhadap reputasi artis atau institusi hiburan. Salah satu contohnya adalah fenomena masifnya sentimen negatif yang dialami oleh Liga Sepak Bola Saudi akibat perilaku pemain asing yang dianggap tidak sejalan dengan nilai budaya setempat [1]. Kasus ini menunjukkan betapa besar kekuatan opini publik dalam membentuk persepsi dan citra suatu institusi.

Di sisi lain, gelombang budaya populer Korea (Korean Wave) di Indonesia menunjukkan bahwa euforia penggemar dapat menjadi kekuatan besar dalam membentuk opini publik terhadap budaya asing, meskipun tidak jarang juga memicu penolakan jika dianggap mengancam nilai budaya lokal [2]. Kondisi ini menunjukkan betapa kompleksnya dinamika sosial di balik budaya populer saat ini. Bahkan, muncul berbagai dilema emosional di kalangan penggemar, seperti saat idola mereka terlibat dalam kampanye produk yang mendapat boikot publik. Situasi seperti ini kerap memunculkan konflik batin antara loyalitas terhadap idola dan komitmen terhadap prinsip sosial atau keagamaan [3]. Tak kalah menarik, kehadiran figur hiburan berbasis kecerdasan buatan seperti penyanyi virtual hasil pengembangan AI turut memicu perdebatan publik. Masyarakat mempertanyakan keaslian, etika, hingga dampaknya terhadap ekosistem hiburan konvensional [4]. Fenomena-fenomena ini menunjukkan bahwa memahami opini publik bukan sekadar kebutuhan, melainkan menjadi strategi penting dalam pengelolaan reputasi, perumusan kebijakan pemasaran, hingga mitigasi krisis di industri hiburan global [1], [4].

Selain itu, ikatan emosional yang terbentuk antara fans dan idola dapat mendorong perilaku konsumtif yang bersifat simbolik dan afektif [5]. Fenomena serupa juga ditemukan pada penggemar JKT48 di Indonesia, di mana perilaku konsumtif didorong oleh fanatisme dan kebutuhan akan kedekatan emosional dengan idola [6]. Di era media sosial, hubungan fans-idola mengalami pergeseran menjadi interaksi parasosial yang lebih intens, personal, dan emosional [7]. Dalam komunitas ini, merchandise tidak lagi sekadar produk, melainkan simbol loyalitas, identitas sosial, dan bentuk afeksi terhadap idola [8], [9].

Oleh karena itu, respons terhadap perilisan merchandise di media sosial dapat mencerminkan lebih sekedar opini terhadap produk, melainkan juga dinamika hubungan emosional dan sosial antara fans dan idola. Untuk memahami respons tersebut, analisis sentimen menjadi pendekatan yang relevan. Analisis Sentimen merupakan teknik dalam bidang Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi opini berbasis teks dan mengelompokkannya ke dalam kategori seperti positif, negatif, dan netral [10]. Salah satu algoritma yang sering digunakan dalam analisis adalah Support Vector Machine (SVM), yang efektif dalam mengolah data berdimensi tinggi dan cocok digunakan untuk klasifikasi teks [11]. Oleh sebab itu, penerapan analisis sentimen menjadi kebutuhan mendesak, tidak hanya untuk memetakan persepsi publik, tetapi juga untuk membantu industri hiburan dalam mengambil keputusan strategis di tengah dinamika sosial yang semakin kompleks.

Meskipun berbagai penelitian terkait analisis sentimen di media sosial telah dilakukan, sebagian besar masih berfokus pada isu-isu umum seperti politik, layanan publik, atau produk komersial. Studi yang secara khusus mengkaji respons publik terhadap perilisan merchandise oleh komunitas penggemar idola di Indonesia, terutama dalam konteks grup JKT48, masih sangat terbatas. Selain itu, beberapa penelitian terdahulu umumnya

menggunakan metode klasifikasi seperti Naïve Bayes atau Long Short-Term Memory (LSTM). Namun, metode tersebut memiliki keterbatasan dalam menangani data berdimensi tinggi dan pola data yang kompleks. Oleh karena itu, penelitian ini menggunakan algoritma Support Vector Machine (SVM) yang dikenal efektif untuk tugas klasifikasi teks, serta mengintegrasikan teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk mengatasi ketidakseimbangan data, sebagaimana juga telah terbukti mampu meningkatkan akurasi model dalam penelitian sebelumnya [12]. Dengan kombinasi tersebut, diharapkan penelitian ini dapat memberikan kontribusi dalam pengembangan penerapan Natural Language Processing (NLP) di bidang analisis sentimen.

Tujuan dari penelitian ini adalah untuk menerapkan analisis sentimen menggunakan algoritma Support Vector Machine (SVM) terhadap respons publik di media sosial, khususnya terkait perilsan merchandise oleh grup idola, guna memahami kecenderungan opini serta dinamika sosial-emosional yang melatarbelakanginya.

2. METODE

Penelitian ini menggunakan pendekatan text mining, yaitu proses pengolahan data teks tidak terstruktur menjadi informasi yang dapat dianalisis secara sistematis [13]. Proses text mining mencakup beberapa tahapan, mulai dari pengumpulan data dan pemrosesan data hingga konversi teks ke bentuk numerik dan analisis menggunakan pendekatan statistik maupun algoritma machine learning. Data yang digunakan berupa komentar dari platform X terkait perilsan merchandise ulang tahun oleh anggota grup idola JKT48. Komentar-komentar ini dikumpulkan menggunakan Tweet Harvest, yang menghasilkan 1.062 komentar. Alur lengkap proses penelitian ditampilkan pada Gambar 1.

Alur Kerja Klasifikasi dengan SVM



Gambar 1. Alur Penelitian

Setelah data terkumpul, dilakukan pelabelan sentimen secara manual ke dalam tiga kategori: positif, negatif, dan netral. Proses pelabelan mempertimbangkan makna dan konteks dalam setiap komentar. Selanjutnya, data melalui tahap pre-processing (seperti case folding, cleansing, tokenisasi, dan stopword), kemudian dikonversi menjadi vector fitur menggunakan metode TF-IDF. Data yang telah direpresentasikan ini digunakan untuk melatih model klasifikasi menggunakan algoritma Support Vector Machine (SVM). Untuk mengatasi ketidakseimbangan kelas, diterapkan teknik Synthetic Minority Over-sampling Technique (SMOTE). Evaluasi performa model dilakukan menggunakan metrik akurasi dan F1-score, serta validasi dengan 5-fold cross-validation.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data dalam penelitian ini berasal dari komentar pengguna pada beberapa thread unggahan resmi anggota JKT48 di platform X yang secara eksplisit mempromosikan merchandise ulang tahun mereka yaitu Birthday T-Shirt. Pengumpulan data dilakukan menggunakan Tweet Harvest, alat web scraping berbasis Node.js untuk mengambil komentar (reply). Total data yang berhasil dikumpulkan sebanyak 1.062 komentar. Komentar-komentar ini menjadi sumber utama dalam proses analisis sentimen, mulai dari tahap pelabelan, pre-processing, hingga pelatihan dan evaluasi model.

3.2. Labeling

Setelah data dikumpulkan, dilakukan proses pelabelan komentar secara manual ke dalam tiga kategori sentimen: positif, negatif dan netral. Penentuan label didasarkan pada makna komentar yang ditulis oleh pengguna, dengan memperhatikan konteks, gaya bahasa, dan ekspresi emosional.

Komentar positif mencerminkan dukungan atau antusiasme terhadap merchandise, netral bersifat informatif tanpa emosi tertentu, dan negatif mengandung ketidaktertarikan atau keberatan, meskipun tidak bersifat ofensif. Pelabelan dilakukan oleh peneliti berdasarkan pemahaman terhadap isi komentar dalam konteks thread masing-masing anggota.

Frequency-Inverse Document Frequeuncy (TF-IDF) [20]. TF-IDF memberikan bobot pada kata berdasarkan frekuensi kemunculannya dalam suatu dokumen dan seberapa jarang kata tersebut muncul di seluruh dokumen [21].

3.5. Evaluasi Model

Model Support Vector Machine (SVM) pertama kali dilatih menggunakan pendekatan train-test split dengan rasio 80:20. Hasil evaluasi model ditunjukkan pada Tabel 5.

Tabel 5. Hasil Evaluasi Model

Label	Precision	Recall	F1-Score	Support
Negatif	0.8750	0.4375	0.5833	16
Netral	0.8448	0.9074	0.8750	108
Positif	0.8452	0.8452	0.8452	84
Accuracy			0.8462	208
Macro Avg	0.8550	0.7300	0.7679	
Weighted Avg	0.8473	0.8462	0.8405	

Dari hasil tersebut, terlihat bahwa performa model terhadap kelas positif dan netral cukup baik. Namun, performa untuk kelas negatif masih rendah, terutama pada metrik recall, yaitu hanya 0.4375. Hal ini disebabkan oleh ketidakseimbangan data pada masing-masing label dimana komentar negatif merupakan kelas dengan jumlah paling sedikit. Untuk mengatasi masalah tersebut, diterapkan metode Synthetic Minority Over-sampling Technique (SMOTE). SMOTE bekerja dengan cara menambahkan data baru secara otomatis pada kelas yang jumlahnya lebih sedikit [22]. Data ini dibuat di antara sampel yang mirip, sehingga jumlah data tiap kelas menjadi seimbang dan model bisa belajar lebih adil [23]. Hasil evaluasi model setelah dilakukan oversampling menggunakan SMOTE ditunjukkan pada Tabel 6.

Tabel 6. Hasil dengan SMOTE

Label	Precision	Recall	F1-Score	Support
Negatif	0.9352	0.9439	0.9395	107
Netral	0.8348	0.8889	0.8610	108
Positif	0.9400	0.8704	0.9038	108
Accuracy			0.9009	323
Macro Avg	0.9033	0.9011	0.9015	
Weighted Avg	0.9032	0.9009	0.9013	

Hasil tersebut menunjukkan peningkatan yang signifikan, khususnya pada kelas negatif yang sebelumnya memiliki nilai recall rendah. Setelah dilakukan SMOTE, nilai recall naik menjadi 0.9439, dan f1-score juga meningkat menjadi 0.9359

Untuk mengevaluasi kestabilan model, dilakukan pengujian menggunakan teknik 5-Fold Cross Validation. Hasil perbandingan dapat dilihat pada Tabel 7.

Tabel 7. Hasil Validasi Model

Metode	Akurasi	F1-Score
Tanpa SMOTE	0.8359	0.7453
Dengan SMOTE	0.9088	0.9088

Hasil uji performa model dengan teknik 5-fold cross validation menunjukkan bahwa model SVM tanpa penerapan SMOTE menghasilkan rata-rata akurasi sebesar 0.8359 dan F1-score sebesar 0.7453. Setelah diterapkan teknik SMOTE untuk menyeimbangkan distribusi kelas, performa model meningkat secara signifikan dengan akurasi dan F1-score yang sama mencapai 0.9088

Peningkatan ini menunjukkan bahwa SMOTE efektif dalam meningkatkan generalisasi model, terutama dalam mengenali pola dari kelas minoritas yang sebelumnya kurang terwakili. Dengan demikian, penambahan SMOTE berkontribusi positif dalam meningkatkan akurasi dan kesimbangan klasifikasi sentimen.

4. KESIMPULAN

Berdasarkan hasil pengujian, model SVM tanpa penerapan SMOTE menghasilkan akurasi sebesar 0.8462 dan F1-score sebesar 0.7679. Setelah menerapkan teknik SMOTE, performa model meningkat dengan akurasi mencapai 0.9009 dan F1-score sebesar 0.9015. Hasil dengan 5-fold cross validation juga menunjukkan peningkatan, di mana akurasi naik dari 0.8359 menjadi 0.9088, dan F1-score dari 0.7453 menjadi 0.9088.

Dengan demikian, dapat disimpulkan bahwa penerapan SMOTE memberikan dampak positif terhadap performa model, khususnya dalam meningkatkan kemampuan klasifikasi terhadap kelas minoritas

REFERENSI

- [1] D. A. Q. Alraimi and I. V. Papatungan, "Decoding Fan and Societal Sentiment: ABSA of The Saudi Pro League's Recent Evolution," *J. Sains, Nalar, dan Apl. Teknol. Inf.*, vol. 4, no. 1, pp. 44–54, Jan. 2025, doi: 10.20885/snati.v4.i1.6.
- [2] P. R. A. Savitri, I. M. A. D. Suarjaya, and W. O. Vihikan, "Sentiment Analysis of X (Twitter) Comments on The Influence of South Korean Culture in Indonesia," *J. Inf. Syst. Informatics*, vol. 6, no. 2, pp. 979–991, 2024, doi: 10.51519/journalisi.v6i2.749.
- [3] K. Fitria, "Love vs. principle: Fans' dilemma over idols promoting or collaborating with boycotted products," *J. Komun.*, vol. 19, no. 2, pp. 59–74, Apr. 2025, doi: 10.20885/komunikasi.vol19.iss2.art4.
- [4] S. Il Shin, J. S. Kim, J. B. Kim, O. K. D. Lee, and H. Yun, "What Do You Feel about an AI Singer Created by Generative AI? Exploring Sentiment about an AI Singer from a Dual Realism Perspective," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 8, pp. 503–512, 2025.
- [5] L. Chen, G. Chen, S. Ma, and S. Wang, "Idol Worship: How Does It Influence Fan Consumers' Brand Loyalty?," *Front. Psychol.*, vol. 13, no. April, pp. 1–14, 2022, doi: 10.3389/fpsyg.2022.850670.
- [6] G. Damasta and D. Dewi, "Hubungan Antara Fanatisme Dengan Perilaku Konsumtif Pada Fans JKT48 di Surabaya," *J. Penelit. Psikol.*, vol. 7, no. 4, pp. 13–18, 2020.
- [7] F. Yasmin, "Konsumerisme Pada Penggemar K-Pop di Era Ekonomi Digital Berbasis Autbase Twitter," *J. Diskurs. Pendidik. Sosiol.*, vol. 5, no. 01 SE-Articles, pp. 22–43, Jun. 2024, [Online]. Available: <https://journal.unj.ac.id/unj/index.php/jdps/article/view/46653>
- [8] A. Maulidia and M. F. Aransyah, "Pengaruh Kepercayaan dan Harga terhadap Keputusan Pembelian dan Loyalitas Pelanggan pada pembelian Merchandise K-Pop melalui Group Order," *J-MAS (Jurnal Manaj. dan Sains)*, vol. 8, no. 1, p. 1022, 2023, doi: 10.33087/jmas.v8i1.1026.
- [9] P. Ciptaningtyas and A. R. Prasetyo, "Nomor 06," *J. Empati*, vol. 11, pp. 381–385, 2022.
- [10] E. Damayanti, Junaedy, and H. B., "Analisis sentimen penggemar grup K-pop NCT pada media sosial X (Twitter) menggunakan algoritma Support Vector Machine (SVM)," *J. Teknol. dan Komput.*, vol. 4, no. 02 SE-Articles, pp. 483–490, Dec. 2024, doi: 10.56923/jtek.v4i02.188.
- [11] N. Maulidah, R. Supriyadi, D. Y. Utami, F. N. Hasan, A. Fauzi, and A. Christian, "Prediksi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine dan Naive Bayes," *Indones. J. Softw. Eng.*, vol. 7, no. 1, pp. 63–68, 2021, doi: 10.31294/ijse.v7i1.10279.
- [12] F. Suandi *et al.*, "Enhancing Sentiment Analysis Performance Using SMOTE and Majority Voting in Machine Learning Algorithms," in *Proceedings of the 7th International Conference on Applied Engineering (ICAIE) 2024*, Batam, Indonesia: Atlantis Press, 2024, pp. 126–138. doi: 10.2991/978-94-6463-620-8_10.
- [13] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, "Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.
- [14] F. F. Irfani, "Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine," *JBMI (Jurnal Bisnis, Manajemen, dan Inform.)*, vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.
- [15] T. M. Permata Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19," *SINTECH (Science Inf. Technol. J.)*, vol. 4, no. 2, pp. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [16] 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.
- [17] R. C. Rivaldi, T. D. Wismarini, J. T. Lomba, and J. Semarang, "Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee)," *ELKOM*, vol. 17, no. 1, pp. 120–128, 2024, doi: 10.51903/elkom.v17i1.1680.
- [18] M. R. W. Julianto, D. Wasistha, and A. D. Hartanto, "Implementasi Algoritma Support Vector Machine Untuk Sentimen Analisis Cyberbullying Pada Kolom Komentar Instagram Artis," *Intechno J. (Information Technol. Journal)*, vol. 3, no. 2, pp. 27–31, 2021, doi: 10.24076/intechnojournal.2021v3i2.1553.
- [19] E. N. Hamdana, A. Okta, N. Wardani, A. Retno, and T. Hayati, "Sentiment Analysis of Visitor Reviews on Google Maps at Kampung Coklat Tourism," vol. 5, no. 1, pp. 274–282, 2025, doi: 10.30811/jaise.v5i1.6488.
- [20] T. Hidayat, N. Nurchim, and N. Nurmalitasari, "Analisis Sentimen Ulasan Pengguna pada Aplikasi SIREKAP 2024 Menggunakan Machine Learning," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 14, no. 1, pp. 109–116, Jan. 2025, doi: 10.30591/smartcomp.v14i1.7021.
- [21] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks," *SINTESIA J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 2, pp. 81–88, 2023.
- [22] H. H. Limbong and N. N. Norhikmah, "Optimization of Sentiment Analysis for Amikom One Application Reviews Using SMOTE with Artificial Neural Network Algorithm," *SISTEMASI*, vol. 13, no. 5, p. 2048, Sep. 2024, doi: 10.32520/stmsi.v13i5.4437.
- [23] Robert Antonius, A. R. Zulkarnain, and H. Irsyad, "Pendekatan TF-IDF, SMOTE, dan SVM dalam Klasifikasi Sentimen Masyarakat terhadap Pemblokiran Judi Online," *Bul. Ilm. Inform. Teknol.*, vol. 2, no. 3, pp. 115–122, 2024, doi: 10.58369/biit.v2i3.65.