

Analisis Klasifikasi Tweet Berdasarkan Topik Sosial Menggunakan SVM

Abdurrazik^{a1}, I Made Widhi Wirawan^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹razikajee@gmail.com
²made_widhi@unud.ac.id

Abstract

Social media platforms, including Twitter (now X), produce a constant flow of user-generated text that reflects public discourse in real time. However, the informal and unstructured nature of these short messages poses challenges for manual topic classification, especially when handling large volumes. This study aims to categorize Indonesian-language tweets into three topics: Politics, Entertainment, and Others, using a supervised machine learning approach. A total of 1,478 tweets were collected through keyword-based scraping and manually labeled according to predefined guidelines. The preprocessing stage included text cleaning, tokenization, stopword removal, stemming, and label encoding. TF-IDF was employed to convert the cleaned text into numerical features, while classification was performed using the Support Vector Machine (SVM) algorithm with a One-vs-Rest strategy for multi-class classification. The model reached an overall accuracy of 84 percent, with particularly high performance in the Politics and Entertainment categories. These results indicate that the combination of TF-IDF and SVM is effective for classifying short Indonesian-language tweets and can be applied to support the organization and filtering of topical content in social media analytics.

Keywords: Twitter, Topic Classification, TF-IDF, Support Vector Machine

1. Pendahuluan

Media sosial saat ini memainkan peran sentral dalam kehidupan masyarakat dan menjadi salah satu sumber informasi utama, baik bagi individu maupun institusi dengan kepentingan politik atau ekonomi. Di antara berbagai platform yang ada, X (Twitter) menonjol sebagai platform *microblogging* yang sangat populer, dengan jutaan pengguna aktif bulanan di seluruh dunia, termasuk di Indonesia [1], [2]. Pengguna memanfaatkan X (Twitter) untuk berbagi pemikiran, sentimen, dan pembaruan status atau "tweet" tentang berbagai topik dan peristiwa secara *real-time*. Meskipun data tekstual mentah yang dihasilkan oleh pengguna X (Twitter) seringkali bersifat semi-terstruktur, bising (*noisy*), dan menggunakan berbagai terminologi atau *slang* yang tidak terstruktur dengan baik, data ini tetap mengandung informasi penting dan detail yang berharga. Mengingat volume informasi yang dihasilkan oleh X (Twitter) sangatlah besar, mencapai ribuan *tweet* per detik, proses pemindaian manual terhadap seluruh dataset menjadi sulit dan tidak efisien [3], [4]. Oleh karena itu, ekstraksi informasi tersembunyi ini menjadi sangat krusial untuk pengambilan keputusan yang tepat dan penyusunan rencana strategis, seperti dalam memahami umpan balik atau memasuki pasar baru [5]. Untuk mengekstrak informasi berharga dari data bahasa ini, teknik Pemrosesan Bahasa Alami (*Natural Language Processing/NLP*) dan penambangan data telah menjadi bidang studi yang sangat menonjol. Bidang ini secara khusus diaplikasikan dalam konteks analisis sentimen dan klasifikasi teks. Analisis sentimen merupakan proses mengidentifikasi perasaan, opini, dan wawasan yang tersembunyi dalam data, kemudian mengkategorikannya berdasarkan polaritas seperti positif, negatif, atau netral. Sementara itu, klasifikasi teks, yang juga dikenal sebagai penandaan atau pelabelan teks, melibatkan pengorganisasian teks berurutan ke dalam kategori-kategori yang berbeda berdasarkan kontennya. Aplikasi NLP dan penambangan data ini sangat penting bagi bisnis, penyedia layanan, dan pembuat kebijakan untuk mengevaluasi kembali pendekatan mereka, meningkatkan penawaran, memahami opini publik, mendeteksi tren yang muncul, dan

membimbing proses pengambilan keputusan di era digital [1], [5].

Namun, di platform media sosial seperti X (Twitter), *tweet* berita atau informasi seringkali tidak dikategorikan secara otomatis berdasarkan kategori bahasan spesifiknya, misalnya berita politik, atau hiburan. Ketiadaan kategorisasi otomatis ini menyulitkan pengguna untuk menemukan informasi sesuai kategori yang mereka inginkan, memaksa mereka untuk melakukan *scroll* panjang pada akun media *online* atau mencari berdasarkan kata kunci judul, yang tentunya sangat merepotkan. Mengingat bahwa klasifikasi berita dapat menjadi proses yang padat karya dan membutuhkan upaya manual yang signifikan, terdapat kebutuhan yang berkembang untuk mengotomatiskan pengelompokan *tweet* berdasarkan topik [2], [5], [6]. Untuk mengatasi tantangan klasifikasi ini, algoritma *Support Vector Machine* (SVM) telah terbukti menjadi salah satu algoritma *machine learning* (ML) berbasis *supervised learning* yang paling efektif dan banyak diaplikasikan. Prinsip kerja SVM didasarkan pada konsep Structural Risk Minimization (SRM), yang berusaha mencari hyperplane optimal untuk memisahkan data ke dalam dua atau lebih kelas dalam ruang fitur berdimensi tinggi. Metode ini dikenal karena kemampuannya yang kuat dalam klasifikasi teks dan regresi, serta keunggulannya dalam menangani dataset yang lebih kecil dan lebih “bersih”. Untuk merepresentasikan data tekstual yang akan diklasifikasikan oleh SVM, TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah teknik pembobotan istilah yang dikenal efektif dan luas digunakan dalam ekstraksi fitur, termasuk untuk mencapai klasifikasi yang akurat [1], [3], [6]. Dalam penelitian ini, data primer dikumpulkan melalui *scraping* dari platform X (Twitter) menggunakan berbagai kata kunci yang relevan dengan topik sosial yang digunakan. Penelitian sebelumnya menunjukkan bahwa SVM dapat digunakan secara optimal dalam berbagai tugas klasifikasi *tweet*, termasuk klasifikasi jenis berita, deteksi sikap (*stance detection*), dan identifikasi perilaku terorganisir, seringkali memberikan akurasi yang tinggi [2], [3], [4], [6], [7]. Dengan mempertimbangkan hal ini, penelitian ini berfokus pada analisis klasifikasi *tweet* berdasarkan topik sosial (Politik, Hiburan, dan Lainnya) menggunakan algoritma SVM, dengan representasi fitur berbasis TF-IDF, bertujuan untuk mengklasifikasikan *tweet* secara otomatis berdasarkan topik utama agar dapat mendukung analisis konten dan pengelompokan informasi secara lebih efisien dan terstruktur. Penelitian ini dilakukan pada hampir 1.500 *tweet* berbahasa Indonesia yang dikumpulkan selama periode Desember 2024 hingga Juni 2025.

2. Metode Penelitian

2.1 Pengumpulan Data

Data yang digunakan diperoleh melalui proses pengumpulan data secara otomatis menggunakan metode *web scraping* dari platform X (Twitter). Proses *scraping* dilakukan terhadap *tweet* berbahasa Indonesia dengan memanfaatkan kata kunci tertentu yang merepresentasikan tiga topik utama, yaitu politik, hiburan, dan lainnya. Kata kunci ini dirancang untuk meningkatkan kemungkinan keberagaman topik dalam data yang dikumpulkan [6]. Jumlah data yang berhasil dikumpulkan sebanyak 1.478 baris *tweet* dalam format teks mentah, masing-masing disimpan dalam format CSV dengan kolom utama berisi isi *tweet* (*full_text*) dan label topik (label).

Data kemudian diklasifikasi secara manual ke dalam tiga label:

- a. Politik: *tweet* yang membahas isu pemerintahan, partai politik, tokoh politik, kebijakan publik, atau peristiwa politik.
- b. Hiburan: *tweet* yang berkaitan dengan musik, film, selebriti, acara budaya pop, olahraga, atau fandom.
- c. Lainnya: *tweet* yang tidak berkaitan dengan dua kategori di atas, seperti opini umum, promosi, informasi netral, atau interaksi harian.

Proses pelabelan dilakukan secara manual oleh satu *annotator* berdasarkan *guideline* pelabelan yang telah disusun sebelumnya untuk menjaga konsistensi dan objektivitas klasifikasi. Setelah proses labeling selesai, distribusi dalam dataset adalah sebagai berikut: Politik (568 data), Hiburan (439 data), dan Lainnya (471 data). Dataset selanjutnya dipisahkan ke dalam dua *subset*, yaitu data pelatihan dan data pengujian, dengan rasio pembagian sebesar 80% dan 20%. Pemisahan dilakukan menggunakan metode *stratified split* guna memastikan distribusi label tetap

proporsional pada masing-masing *subset*.

2.2 *Preprocessing*

Tahapan *preprocessing* dilakukan untuk membersihkan dan menyiapkan data teks agar dapat direpresentasikan secara numerik dan digunakan dalam proses pelatihan model klasifikasi [5], [6]. Proses ini melibatkan beberapa langkah sebagai berikut:

a. **Pembersihan Teks**

Pada tahap ini, elemen-elemen yang tidak memiliki kontribusi semantik terhadap isi tweet dihapus, seperti:

- *Mention* (@username)
- Tautan *URL*
- *Hashtag*
- Karakter khusus dan emoji
- Angka yang tidak relevan

Langkah ini penting untuk mengurangi noise dalam data.

b. **Case Folding**

Seluruh teks diubah menjadi huruf kecil (*lowercase*). Hal ini dilakukan untuk menghindari perbedaan makna antara kata yang sama namun ditulis dengan huruf besar dan kecil.

c. **Tokenisasi**

Proses ini memecah teks menjadi unit-unit kata atau token. Token adalah representasi kata yang akan digunakan dalam tahap ekstraksi fitur selanjutnya.

d. **Stopword Removal**

Stopword adalah kata-kata umum yang tidak memiliki makna penting dalam proses klasifikasi (contohnya: “yang”, “di”, “dan”). Kata-kata ini dihapus untuk mengurangi dimensi fitur dan meningkatkan efisiensi.

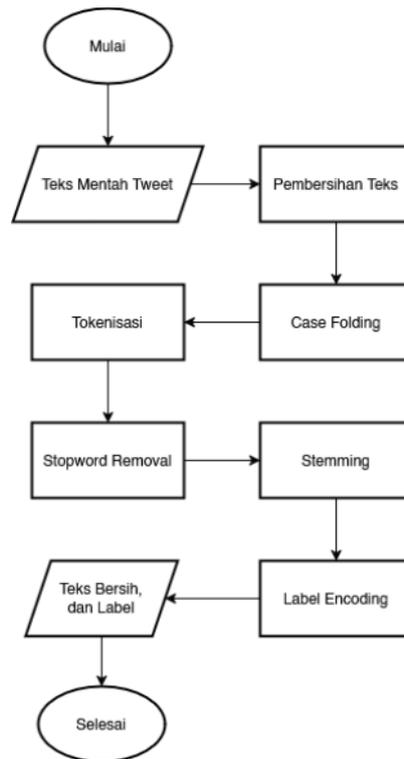
e. **Stemming**

Tahap ini mengubah kata ke bentuk dasarnya. Misalnya, “menonton” dan “penonton” akan dikembalikan ke bentuk dasar “tonton”. Tujuannya adalah menyatukan variasi morfologis dari kata agar model tidak menganggapnya sebagai kata berbeda.

f. **Label Encoding**

Label teks (seperti “politik”, “hiburan”, dan “lainnya”) diubah menjadi angka menggunakan metode encoding. Ini diperlukan agar algoritma klasifikasi dapat memproses label dalam bentuk numerik.

Gambar 1 menunjukkan alur *preprocessing* yang dilakukan terhadap teks *tweet* sebelum masuk ke tahap ekstraksi fitur.



Gambar 1. Flowchart Tahapan Preprocessing

2.3 Term Frequency Inverse Document Frequency (TF-IDF)

Sebelum teks dapat diproses oleh model klasifikasi, diperlukan proses transformasi dari teks mentah menjadi representasi numerik. Pada penelitian ini, representasi dilakukan menggunakan pendekatan *Term Frequency Inverse Document Frequency* (TF-IDF). Pendekatan ini digunakan untuk memberikan bobot pada kata berdasarkan tingkat kepentingannya dalam sebuah dokumen dibandingkan dengan keseluruhan kumpulan dokumen (*corpus*).

TF-IDF terdiri dari dua komponen utama :

- a. *Term Frequency* (TF): Mengindikasikan tingkat kemunculan kata t dalam sebuah dokumen d , dibandingkan dengan total kata di dalam dokumen tersebut.

$$TF(t, d) = \frac{f_{t,d}}{\sum_k f_{k,d}} \quad (1)$$

Keterangan:

$f_{t,d}$ jumlah kemunculan kata t dalam dokumen d
 $\sum_k f_{k,d}$ total seluruh frekuensi semua kata dalam dokumen d

- b. *Inverse Document Frequency* (IDF): Menggambarkan seberapa jarang atau spesifik suatu kata t muncul di seluruh dokumen dalam korpus.

$$IDF(t) = \log\left(\frac{N}{df_t}\right) \quad (2)$$

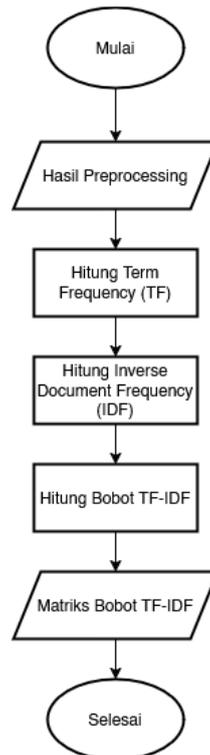
Keterangan:

N adalah jumlah total dari dokumen.
 df_t adalah jumlah dokumen yang mengandung kata t .

c. TF-IDF Score: Hasil akhir dari proses adalah skor gabungan:

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

Dalam studi ini, metode TF-IDF dimanfaatkan untuk merepresentasikan teks tweet ke dalam bentuk vektor numerik dengan dimensi tetap, yang selanjutnya digunakan sebagai masukan untuk proses klasifikasi. Teknik ini terbukti efektif dalam berbagai studi klasifikasi teks media sosial [5], [6]. Jumlah fitur dibatasi hingga 5.000 kata yang paling sering muncul untuk mengurangi kompleksitas dan overfitting. Gambar 2 menjelaskan tahapan transformasi teks bersih menjadi representasi numerik menggunakan metode TF-IDF.



Gambar 2. Flowchart Ekstraksi Fitur TF-IDF

2.4 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan metode klasifikasi yang bersifat diskriminatif, dengan mekanisme utama mencari hyperplane yang paling optimal sebagai batas pemisah antara dua kelas data dalam ruang fitur berdimensi tinggi [6]. SVM sangat sesuai untuk data teks karena fitur hasil ekstraksi seperti TF-IDF menghasilkan vektor berdimensi besar jarang (*sparse*).

Secara umum, SVM mencoba membentuk *hyperplane* dalam bentuk:

$$w \cdot x + b = 0 \quad (4)$$

Dimana x adalah vektor dari tweet, w adalah bobot, dan b adalah bias. Model dilatih untuk menemukan kombinasi w dan b yang dapat memisahkan tweet dari kelas yang berbeda dengan margin selebar mungkin. Tujuan ini dinyatakan dalam bentuk fungsi optimasi:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (5)$$

Untuk data yang tidak sepenuhnya dapat dipisahkan secara linear, digunakan versi *soft-margin* dengan pelanggaran menggunakan *slack variable* ξ_i :

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (6)$$

Keterangan:

x adalah vektor fitur dari data (TF-IDF)

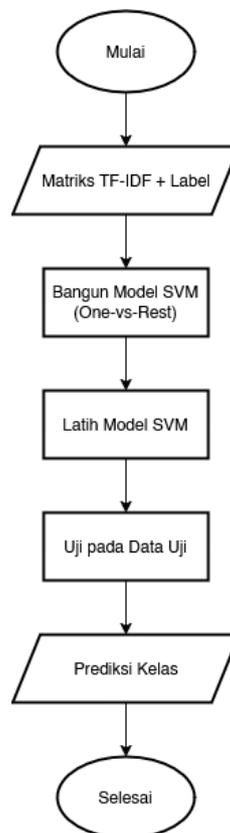
w adalah bobot vektor

b adalah bias (intersep dari *hyperplane*)

C adalah parameter untuk mengontrol seberapa besar model mengizinkan kesalahan

ξ_i adalah nilai toleransi *error* dari masing-masing data

Karena penelitian ini melibatkan tiga kelas (Politik, Hiburan, dan Lainnya), maka digunakan pendekatan *One-vs-Rest* (OVR). Dalam pendekatan ini, model SVM biner dilatih untuk masing-masing kelas terhadap gabungan dua kelas lainnya. Prediksi dilakukan dengan memilih model yang memberikan skor paling tinggi. Kombinasi SVM dan TF-IDF telah banyak digunakan dalam klasifikasi *tweet* dan terbukti memberikan hasil yang kompetitif [4], [5]. Gambar 3 menggambarkan proses pelatihan dan pengujian model klasifikasi menggunakan algoritma SVM.



Gambar 3. Flowchart Klasifikasi dengan SVM

2.5 Evaluasi

Pada tahap evaluasi, *confusion matrix* dimanfaatkan untuk memperoleh metrik performa seperti akurasi, *precision*, dan *F1-score*. Matriks ini merepresentasikan hasil prediksi model klasifikasi dan memberikan gambaran seberapa tepat model dalam menempatkan data ke dalam kelas yang sesuai. Pada kasus multi-kelas, *confusion matrix* diinterpretasikan untuk setiap kelas menggunakan pendekatan *One-vs-Rest*. Untuk dua kelas, bentuk umum *confusion matrix* ditunjukkan pada Tabel 1.

Tabel 1. Struktur *Confusion Matrix* (Dua Kelas)

Kelas Sebenarnya	Prediksi Positif	Prediksi Negatif
Positif	TP	FN
Negatif	FP	TN

Keterangan:

TP = *True Positive* (jumlah prediksi yang benar untuk kelas positif)

FN = *False Negative* (jumlah data positif yang salah diprediksi sebagai negatif)

TN = *True Negative* (jumlah prediksi yang benar untuk kelas negatif)

FP = *False Positive* (jumlah data negatif yang salah diprediksi sebagai positif)

Rumus yang digunakan untuk menghitung metrik evaluasi adalah sebagai berikut:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

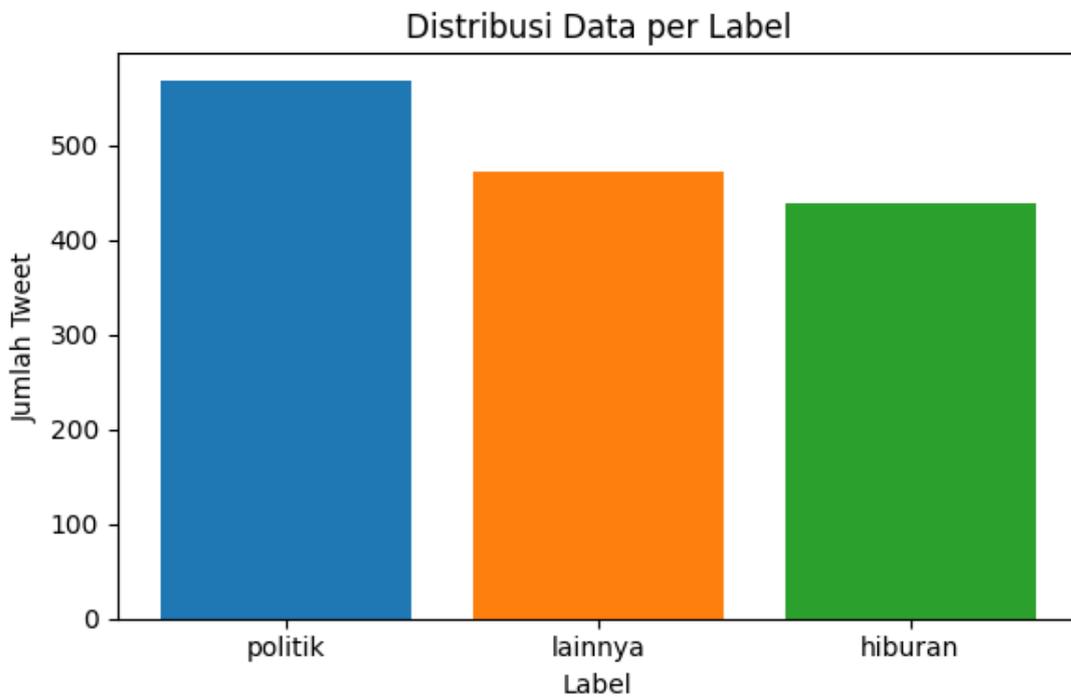
$$F1-Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (9)$$

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

Dalam penelitian ini, evaluasi dilakukan terhadap model klasifikasi tiga kelas, sehingga nilai *precision*, *recall*, dan *F1-score* dihitung masing-masing kelas, lalu dirata-ratakan menggunakan pendekatan *macro average* dan *weighted average* [4].

3. Hasil dan Diskusi

Penelitian ini menghasilkan model klasifikasi topik *tweet* berbahasa Indonesia menggunakan algoritma *Support Vector Machine* (SVM) dengan pendekatan fitur TF-IDF. Model dilatih untuk mengenali tiga kategori topik utama: Politik, Hiburan, dan Lainnya. Hasil evaluasi pada data uji menunjukkan bahwa model mencapai akurasi sebesar 84%, dengan performa terbaik pada kategori Hiburan dan Politik. Temuan ini sejalan dengan hasil dari studi sebelumnya yang menunjukkan bahwa SVM merupakan salah satu metode yang efektif untuk klasifikasi teks pendek dari media sosial, bahkan dengan ukuran dataset yang terbatas [2], [4]. Visualisasi data dari ketiga kelas pada dataset dapat dilihat pada Gambar 4. Kategori Politik merupakan kelas dengan jumlah data terbanyak, diikuti oleh Lainnya dan Hiburan.



Gambar 4. Grafik Distribusi Data per Label

Model SVM diujikan terhadap 20% data uji yang telah dibagi secara *stratified*. Hasil evaluasi ditampilkan dalam bentuk metrik *precision*, *recall*, *F1-score*, serta akurasi total. Hasil evaluasi model pada data uji ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Model SVM

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Hiburan	0,91	0,92	0,92	88
Lainnya	0,89	0,82	0,85	94
Politik	0,87	0,91	0,89	114
Akurasi			0,89	296
Macro avg	0,89	0,88	0,88	296
Weighted avg	0,89	0,89	0,88	296

Untuk memperjelas hasil klasifikasi per kelas, berikut ditampilkan *confusion matrix* dari hasil prediksi terhadap data uji pada Tabel 3.

Tabel 3. *Confusion Matrix*

	Hiburan	Lainnya	Politik
Hiburan	81	4	3
Lainnya	4	77	13
Politik	4	6	104

Dari confusion matrix di atas, terlihat bahwa model memiliki kemampuan yang tinggi dalam

mengklasifikasikan tweet bertopik Hiburan dan Politik secara tepat, ditandai dengan tingginya jumlah prediksi benar pada kedua kelas tersebut. Sementara itu, kelas Lainnya masih menunjukkan tantangan tersendiri, terutama karena kontennya yang lebih beragam dan tidak memiliki indikator topik yang eksplisit. Hal ini menyebabkan sebagian tweet dalam kelas Lainnya terklasifikasi sebagai Politik atau Hiburan. Kesalahan ini menunjukkan bahwa model cenderung kesulitan membedakan konten yang bersifat netral atau ambigu secara topikal, yang memang merupakan karakteristik utama dari kelas Lainnya.

Hasil akurasi keseluruhan sebesar 89% menunjukkan bahwa metode SVM dan TF-IDF cukup efektif dalam menangani klasifikasi topik pada *tweet* pendek berbahasa Indonesia. Nilai *F1-score* yang tinggi pada kelas Hiburan dan Politik mengindikasikan keseimbangan antara *precision* dan *recall* yang baik pada kedua kelas tersebut. Namun demikian, terdapat beberapa tantangan dalam mendeteksi kelas Lainnya, karena topik yang masuk ke dalamnya bersifat sangat beragam dan tidak memiliki pola leksikal yang kuat seperti kelas lainnya. Hal ini berpotensi menurunkan *recall* karena model kesulitan mengenali konten netral atau tidak jelas topiknya. Penurunan performa pada kelas Lainnya sejalan dengan temuan Küçük dan Can, yang mengamati bahwa kategori dengan cakupan topik yang lebih umum cenderung memiliki nilai *recall* yang lebih rendah akibat ambiguitas data [7].

Penggunaan satu *annotator* juga menjadi keterbatasan yang dapat menimbulkan bias subjektif meskipun telah dibantu dengan panduan pelabelan. Untuk penelitian selanjutnya, disarankan untuk menggunakan lebih dari satu *annotator* dan melakukan *inter-annotator agreement* untuk meningkatkan kualitas pelabelan.

4. Kesimpulan

Penelitian ini membahas penerapan algoritma *Support Vector Machine* (SVM) untuk klasifikasi topik *tweet* berbahasa Indonesia menggunakan pendekatan representasi fitur berbasis TF-IDF. Dataset yang digunakan terdiri dari 1.478 *tweet* yang telah dilabeli secara manual ke dalam tiga kategori, yaitu Politik, Hiburan, dan Lainnya. Model SVM dibangun menghasilkan akurasi sebesar 84%, dengan nilai *F1-score* yang cukup tinggi pada seluruh kelas, terutama pada kategori Politik dan Hiburan. Hasil ini menunjukkan bahwa kombinasi antara metode ekstraksi fitur TF-IDF dan algoritma SVM mampu mengklasifikasikan *tweet* pendek dengan efektif, meskipun data bersifat tidak terstruktur dan mengandung variasi bahasa informal [4], [5]. Secara keseluruhan, penelitian ini berhasil menunjukkan bahwa pendekatan berbasis *machine learning* dapat digunakan untuk mengelompokkan informasi di media sosial secara otomatis dan akurat, serta berpotensi membantu dalam pengorganisasian data teks berdasarkan topik sosial tertentu.

Daftar Pustaka

- [1] V. KP, R. AB, G. HL, V. Ravi, dan M. Krichen, "A tweet sentiment classification approach using an ensemble classifier," *International Journal of Cognitive Computing in Engineering*, vol. 5, hlm. 170–177, Jan 2024, doi: 10.1016/j.ijcce.2024.04.001.
- [2] K. E. Yilmaz dan O. Abul, "Inferring Political Alignments of Twitter Users".
- [3] E. Beğenilmiş dan S. Üsküdarlı, "Organized Behavior Classification of Tweet Sets using Supervised Learning Methods," hlm. 1–51, Nov 2017, Diakses: 21 Juni 2025. [Daring]. Tersedia pada: <https://www.fbi.gov/cve508/teen-website/what-are-known-violent-extre>
- [4] S. L. Lo, R. Chiong, dan D. Cornforth, "Using support vector machine ensembles for target audience classification on Twitter," *PLoS One*, vol. 10, no. 4, hlm. 1–20, Apr 2015, doi: 10.1371/journal.pone.0122855.
- [5] M. Hasan, T. Ahmed, M. R. Islam, dan M. P. Uddin, "Leveraging textual information for social media news categorization and sentiment analysis," *PLoS One*, vol. 19, no. 7 JULY, hlm. 1–28, Jul 2024, doi: 10.1371/journal.pone.0307027.
- [6] F. M. Suryana, I. Cholissodin, dan E. Santoso, "Klasifikasi Jenis Berita pada Sosial Media Twitter menggunakan Algoritme Support Vector Machine (SVM)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 6, hlm. 1906–1914, Jun 2020, [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>

- [7] D. Küçük dan F. Can, "Stance Detection on Tweets: An SVM-based Approach," no. 23 Mar, hlm. 1–13, Mar 2018, Diakses: 21 Juni 2025. [Daring]. Tersedia pada: <http://ceng.metu.edu.tr/>