

Analisis Sentimen pada Ulasan Aplikasi Social Media menggunakan Metode Svm dan Decision Tree

Santana Ellyszza, Siti Sa'uda

Jurusan Teknik Informatika, Fakultas Sains Teknologi

Jurusan Teknik Informatika, Fakultas Sains Teknologi

Universitas Bina Darma

santanaellyszza@gmail.com, siti_sauda@binadarma.ac.id

Abstract- In general, social media is used to share information and opinions, including learning about new trends, news, entertainment and other topics. TikTok is a social media platform that focuses on short videos. X/Twitter is a microblogging platform that allows users to send and read short messages, while Instagram is a photo and video sharing platform that allows users to take, edit and share images and short videos. With the existence of social media applications such as TikTok, X/Twitter, and Instagram, this study focuses on these three platforms based on several considerations. First, all three are among the most popular and widely used social media applications in Indonesia and globally, resulting in a large number of user reviews on the Play Store. Second, TikTok, X/Twitter, and Instagram have different characteristics in terms of their main features (short videos, microblogging, and photo/video sharing), allowing for a more diverse sentiment analysis. Third, the selection of these applications is also based on their current relevance, as all three are often in the public and media spotlight, meaning that analysing user perceptions of these applications can provide significant insights. However, both systems have various weaknesses experienced by their users. By evaluating these weaknesses, companies can identify problems in their applications and find ways to improve them. The following research aims to analyse review data on social media applications by utilising two methods, namely Support Vector Machine and Decision Tree. This study will determine which method is more effective and accurate in classifying positive and negative sentiments from review comments. Through classification, the accuracy of Support Vector Machine was found to be 84.61% for the TikTok application, 86.82% for the Twitter, and 85.73% for the Instagram application, while the Decision Tree method achieved 77.98% for the TikTok application, 80.14% for the Twitter application, and 81.71% for the Instagram application. Based on the analysis results, the Support Vector Machine method demonstrated higher accuracy than the Decision Tree method.

Keywords: Social Media, Data Review, Support Vector Machine, Decision Tree

Abstrak- Secara umum social media digunakan untuk berbagi informasi dan pendapat, termasuk mengetahui tren baru, berita, hiburan dan topik lainnya. Tiktok adalah platform media sosial yang fokus pada video pendek. X/Twitter adalah platform mikroblogging yang memungkinkan pengguna mengirim dan membaca pesan singkat, sedangkan Instagram yakni platform berbagi foto dan video yang memungkinkan pengguna mengambil, mengedit, dan membagikan gambar serta video pendek. Namun, kedua sistem ini memiliki berbagai kelebihan yang dialami oleh penggunanya. Dengan adanya aplikasi media sosial seperti TikTok, X/Twitter, dan Instagram, penelitian ini berfokus pada ketiga platform tersebut berdasarkan beberapa pertimbangan. Pertama, ketiganya termasuk dalam kategori aplikasi media sosial yang paling populer dan banyak digunakan di Indonesia maupun secara global, sehingga memiliki jumlah ulasan pengguna yang sangat besar di Play Store. Kedua, TikTok, X/Twitter, dan Instagram memiliki karakteristik yang berbeda dalam hal fitur utama (video pendek, *microblogging*, dan berbagi foto/video) sehingga memungkinkan analisis sentimen yang lebih beragam. Ketiga, pemilihan aplikasi tersebut juga didasarkan pada relevansi aktual, di mana ketiganya sering menjadi sorotan publik dan media, sehingga analisis persepsi pengguna terhadap aplikasi ini dapat memberikan wawasan yang signifikan. Dengan mengevaluasi kelebihan-kelebihan tersebut, perusahaan dapat mengidentifikasi masalah pada aplikasinya dan menemukan cara untuk memperbaikinya. Penelitian berikut bertujuan guna menganalisis data ulasan pada aplikasi media sosial dengan memanfaatkan dua metode, yaitu *Support Vector Machine* dan *Decision Tree*. Penelitian berikut akan menentukan metode mana yang lebih efektif dan akurat mengklasifikasi sentimen positif dan negatif dari review komentar. Melalui klasifikasi diperoleh hasil akurasi *Support Vector Machine* sebesar 84.61% untuk aplikasi Tiktok, 86.82% untuk aplikasi Twitter dan 85.73% untuk aplikasi Instagram, sedangkan metode *Decision Tree* sebesar 77.98% untuk aplikasi Tiktok, 80.14% untuk aplikasi Twitter, dan 81.71% untuk aplikasi Instagram. Berdasarkan hasil analisis, metode *Support Vector Machine* menunjukkan akurasi yang lebih tinggi daripada metode *Decision Tree*.

Kata Kunci: Social Media, Data Ulasan, Support Vector Machine, Decision Tree

1. Pendahuluan

Di zaman globalisasi kini *social media* sudah dijadikan platform populer untuk berbagi informasi dan pendapat,

termasuk mengetahui tren baru, berita, hiburan dan topik lainnya. Media sosial adalah platform yang mana bisa

Vol.16 no.2 | Desember 2025

EXPLORE : ISSN: 2087-2062, Online ISSN: 2686-181X / DOI: <http://dx.doi.org/10.36448/jsit.v16i2.4455>



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

menciptakan banyak bentuk berkomunikasi serta menyebarkan beragam berita informasi guna seluruh kalangan kemasyarakatan. Melalui media sosial, tiap individu bisa melangsungkan komunikasi serta banyaknya berita informasi dengan berbagai kalangan kemasyarakatan.[1]. Dengan adanya aplikasi *social media* seperti tiktok, X/Twitter, dan Instagram memungkinkan pengguna secara instan dapat terhubung dengan teman, keluarga bahkan orang asing di mana pengguna dapat mengungkapkan diri membangun jejaring dan mengakses data dengan mudah. Tujuan penelitian ini adalah mengkaji serta menganalisis data ulasan pada aplikasi media sosial menggunakan dua metode klasifikasi, yaitu Support Vector Machine (SVM) dan Decision Tree. Dalam penelitian ini digunakan algoritma *Support Vector Machine (SVM)* dan *Decision Tree* untuk proses klasifikasi. Perbandingan kedua algoritma ini menjadi penting karena keduanya memiliki karakteristik yang berbeda dan sering digunakan pada berbagai permasalahan data. SVM dikenal memiliki performa yang baik pada data berdimensi tinggi dan mampu menghasilkan hyperplane optimal, sedangkan Decision Tree lebih unggul dalam hal interpretabilitas serta mampu menangani data dengan pola non-linear secara lebih sederhana. Dengan demikian, melakukan perbandingan antara SVM dan Decision Tree dapat memberikan gambaran yang lebih komprehensif mengenai algoritma mana yang lebih sesuai diterapkan pada kasus penelitian ini, baik dari segi akurasi maupun kemudahan interpretasi hasil. Fokus penelitian diarahkan pada tiga aplikasi utama, yaitu TikTok, X/Twitter, dan Instagram. Pemilihan ketiga aplikasi ini didasarkan pada beberapa pertimbangan. Pertama, TikTok, X/Twitter, dan Instagram termasuk dalam platform media sosial dengan jumlah pengguna terbesar serta tingkat interaksi yang tinggi, baik di Indonesia maupun secara global. Kedua, ketiga aplikasi tersebut memiliki karakteristik layanan yang berbeda. TikTok berfokus pada konten video pendek, X/Twitter pada *microblogging*, dan Instagram pada berbagi foto dan video sehingga analisis sentimen terhadapnya dapat memberikan perspektif yang lebih beragam. Ketiga, aplikasi-aplikasi ini secara konsisten menjadi sorotan publik dan media, sehingga hasil penelitian mengenai sentimen pengguna terhadapnya diharapkan memberikan kontribusi yang relevan dalam memahami persepsi masyarakat terhadap media sosial populer.[1]. TikTok adalah platform jejaring sosial yang mana penggunanya dimungkinkan bisa membagikan foto maupun video pengguna berbagi video pendek yang berisi nyanyian atau tarian. TikTok merupakan hasil penggabungan dari dua aplikasi sebelumnya, yaitu Douyin dan Musically.[2]. X/Twitter merupakan aplikasi media sosial dimana semua masyarakat memiliki peluang untuk mengunggah pesan, tulisan, foto di platform twitter.[3]. Twitter ialah platform media sosial yang mana dianggap paling populer. Twitter menjadi layanan yang memungkinkan teman-teman, keluarga, kolega guna melangsungkan komunikasi serta menjaga koneksi lewat pertukaran pesan secara cepat serta mempunyai intesitas yang kerap digunakan[4]. Dan Instagram adalah platform

media sosial yang memungkinkan pengguna untuk membagikan foto dan video menarik kepada orang-orang yang tergabung di Instagram[5].

Komentar atau ulasan adalah pendapat evaluasi yang diberikan oleh individu atau pengguna terhadap suatu produk, layanan, atau pengalaman pengguna terhadap aplikasi yang mereka gunakan. Ulasan-ulasan ini sering kali berisi opini, keluhan, atau puji dari pengguna mengenai pengalaman mereka dengan aplikasi tersebut. Pengguna yang merasa tidak puas dengan produk atau layanan biasanya menuliskan keluhan mereka di kolom komentar pada aplikasi di *Play Store*. Di sisi lain, ada juga pengguna yang senang menuliskan ulasan positifnya tentang suatu produk di kolom komentar aplikasi di situs *Play Store*, terlepas dari apakah pelanggan tersebut menyadari atau tidak, ulasan pelanggan yang ditulis dapat mempengaruhi potensial pelanggan dalam jumlah yang signifikan. Oleh sebab itu diperlukan teknik khusus yaitu analisis sentimen yang dapat secara otomatis mengklasifikasi ulasan berdasarkan hasilnya, baik positif maupun negatif.

Tabel 1 Dataset Karakteristik

Aplikasi	Tiktok	X/Twitter	Instagram
Jumlah	15.258	15.272	15.187
Ulasan			
Periode	Januari-	Januari-	Januari-
Waktu	Juni 2024	Juni 2024	Juni 2024
Bahasa	14.496	14.204	14.276
Indonesia			
Bahasa	762	1.068	911
Asing			
Positif	11.070	7.589	11.546
Negatif	4.188	7.729	3.641

Analisis sentimen mengumpulkan opini, emosi, dan penilaian seseorang yang mana ditulis mengenai suatu subjek dengan menggunakan teknik pemrosesan bahasa alami.[6] Dalam ranah data mining, analisis sentimen adalah cabang yang bertugas mengupas, mengolah, dan menarik informasi dari teks yang dihasilkan oleh individu, organisasi, peristiwa, atau topik tertentu. Intinya, ini tentang menggali informasi dari kumpulan data yang ada. Analisis sentimen merupakan bidang penelitian yang masih cukup baru ketika memproses bahasa alami, ini dimaksudkan guna memprediksi adanya elemen subjektif pada teks serta mengidentifikasi sentimen pada opini tersebut. [7]. Analisis sentimen yakni bagian atas text mining guna mempelajari teknik serta menafsirkan juga melaksanakan pengevaluasian emosional pada suatu teks. Tujuan analisis sentimen adalah untuk mengklasifikasikan sentimen pada teks hingga berkategori positif, negatif, netral. [8]. Tujuan analisis sentimen yakni guna memvalidasi apakah teks tersebut memiliki sentimen positif atau negatif. Dalam konteks ulasan aplikasi, analisis sentimen dapat membantu mengungkapkan bagaimana perasaan



pengguna terhadap aplikasi tersebut seberapa puas mereka dengan fitur-fitur yang disediakan.

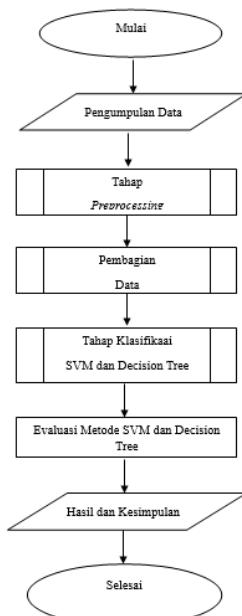
Text mining adalah bagian data mining yang mana memanfaatkan alat analisis guna meraup wawasan atas tumpukan dokumen secara berkala. Konsep ini sering dipakai untuk mengelompokkan dokumen berdasarkan topik dalam proses klasifikasi. Tujuannya adalah membantu pengguna memahami dan menganalisis isi dokumen dengan lebih mudah. [9]. Text mining yakni prosesi pencarian serta pengevaluasian kumpulan data berwujud teks dimaksudkan guna mengekstrak informasi berguna dari data tersebut pada topik tertentu [10].

Tujuan penelitian berikut yakni guna mengkaji serta analisis data ulasan pada aplikasi *social media* menggunakan dua metode yaitu *Support Vector Machine* serta *Decision Tree*. Dalam penerapan algoritma Support Vector Machine (SVM) dan Decision Tree, penelitian ini menggunakan strategi pembagian dataset menjadi data latih dan data uji dengan proporsi 70:30 untuk memastikan perbandingan yang adil. Kedua algoritma dilatih dan diuji pada dataset yang sama sehingga hasil evaluasi dapat dibandingkan secara objektif. Pada SVM digunakan kernel RBF karena memberikan hasil terbaik pada pengujian awal, sedangkan pada Decision Tree dilakukan penyesuaian parameter dasar seperti kedalaman pohon (max depth) untuk mencegah overfitting. Evaluasi kinerja kedua algoritma dilakukan dengan menggunakan metrik akurasi, presisi, recall, F1-score, serta matriks kebingungan. Dengan pendekatan ini, hasil perbandingan dapat mencerminkan performa relatif kedua algoritma tanpa bias dalam pengukuran. Penelitian berikut akan menentukan metode mana yang lebih efektif dan akurat mengklasifikasi sentimen positif maupun negatif dari ulasan. Evaluasi kinerja model dalam penelitian ini tidak hanya didasarkan pada tingkat akurasi, tetapi juga mencakup metrik *precision*, *recall*, dan *F1-score* untuk memperoleh gambaran performa yang lebih komprehensif. Penggunaan beberapa metrik tersebut penting karena dataset ulasan memiliki distribusi kelas yang tidak seimbang antara sentimen positif dan negatif. Dengan mempertimbangkan keempat metrik ini, perbandingan antara algoritma Support Vector Machine dan Decision Tree dapat dilakukan secara lebih objektif dan menggambarkan kemampuan model secara menyeluruh dalam mengklasifikasikan sentimen pengguna. *Support Vector Machine (SVM)* adalah teknik prediksi yang serbaguna, bisa dipakai untuk klasifikasi dan regresi. Awalnya, *SVM* menggunakan prinsip linear classifier untuk menangani kasus klasifikasi yang bisa dipisahkan secara linier. Namun, sekarang *SVM* juga mampu mengatasi masalah non-linier dengan menerapkan konsep kernel di ruang berdimensi tinggi. Di ruang berikut, *SVM* mencari hyperplane yang mana memaksimalkan jarak (*margin*) antara kelas data. Algoritma *Support Vector Machine (SVM)* yakni teknik memprediksi yang sangat efektif untuk berbagai jenis klasifikasi dan regresi. [11]. Sedangkan menurut Mahendra dkk *Support Vectro Machine* adalah metode klasifikasi yang bertujuan memisahkan data ke dalam

beberapa kelas dengan mencari hyperplane optimal yang memaksimalkan jarak antar kelas-kelas tersebut. Algoritma SVM efektif dalam menyelesaikan masalah klasifikasi biner (dua kelas) maupun multikelas.[12]. Akar pohon keputusan (*Decision Tree*) didasarkan dari nilai tertinggi information gain. Metode ini memiliki karakteristik yang paling sesuai dengan objek yang dianalisis. Atribut akar pohon berikut yakni adanya nilai information gain tertinggi, yang menjadi fondasi untuk seluruh struktur pohon keputusan. [13]. Sedangkan menurut Rashid dan Hendra *Decision Tree* yakni sistematis data berisi simpul (*node*) serta rusuk (*edge*). Simpul suatu pohon diklasifikasikan menjadi 3 jenis, yakni simpul akar (*root node*), simpul percabangan (*branch node*), serta simpul daun (*leaf node*). [14]

2. Metodologi

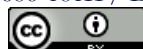
Pada penelitian berikut digunakan, metode Klasifikasi Supervised Learning, sebuah pendekatan yang memanfaatkan data yang sudah ada guna menciptakan pengetahuan termuktahir. Metode ini dilakukan melalui serangkaian langkah seperti yang ditunjukkan gambar berikut.



Gambar 1. Konsep Penelitian

A. Tahap Pengumpulan Data

Tahap berikut, data sekunder dikompilasi muasal berbagai ulasan aplikasi TikTok, X/Twitter, dan Instagram di Play Store. Proses pengumpulan data dilakukan dengan *web scraping* menggunakan source code di Google Colab, yang menghasilkan data sentimen review komentar dalam format CSV. Analisis sentimen yang dilakukan pada penelitian ini tidak difokuskan pada topik tertentu, melainkan bersifat umum untuk mengetahui kecenderungan sentimen pengguna terhadap aplikasi secara keseluruhan positif atau negatif. Pendekatan ini dipilih agar dapat memberikan gambaran umum mengenai persepsi pengguna tanpa membatasi pada aspek tertentu dari aplikasi. Proses pengumpulan



data ulasan aplikasi media Tiktok X/Twiter dan Instagram yang ada pada Play Store dikumpulkan melalui proses *web scraping*. Perolehan Dataset kemudian disimpan berformat CSV. [15]

B. Tahap Preprocessing

Pada tahap ini, data diolah untuk mengekstrak kata yang dibutuhkan dan menghilangkan kata yang tidak relevan untuk analisis melalui beberapa teknik preprocessing, termasuk case folding yang mengubah semua data menjadi huruf kecil, *tokenizing* yang memisahkan teks menjadi kata-kata individual, filtering yang menghapus kata-kata tak penting atau tidak informatif, serta *stemming* yang menghapus imbuhan pada kata sehingga menjadi bentuk kata dasar.

C. Tahap Pembagian Data

Data yang terkumpul dipecah menjadi dua yaitu data latih dan data uji. Data latih diproses menggunakan teknik yang akan digunakan dan hasilnya digunakan untuk memprediksi data uji, sedangkan data uji adalah data yang nantinya dianalisis dan diproyeksikan.

D. Tahap Klasifikasi

Pada tahap ini, terdapat metode yang digunakan untuk klasifikasi, yaitu Support *Vector Machine (SVM)* dan *Decision Tree*. Metode *SVM* menemukan hyperplane paling optimal yang membagi ulasan menjadi dua kelompok, yaitu positif dan negatif, kemudian digunakan untuk prediksi data uji. Sementara itu, *Decision Tree* juga menemukan hyperplane terbaik dengan menggunakan algoritma *Decision Tree* untuk memisahkan ulasan menjadi kelas positif dan negatif, kemudian dipakai untuk proyeksi pada data testing.

E. Tahap Evaluasi SVM dan Decision Tree

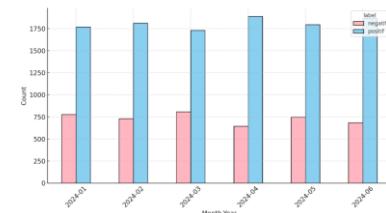
Evaluasi merupakan tahap pengujian metode klasifikasi *SVM* dan *Decision Tree* dengan cara menghitung nilai akurasi masing-masing metode. Tahap berikut, kedua algoritma diuji memanfaatkan *fold= 2* hingga 10 guna mencari skala akurasi optimal. Selanjutnya, kedua metode dikompersasi guna menentukan mana lebih akurat. Jenis *kernel dot, radial, dan polynomial* dipakai untuk *SVM*.

3. Hasil dan Pembahasan

A. Tahap pengumpulan Data

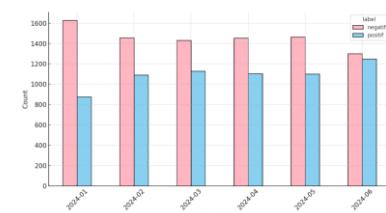
Proses pengumpulan data ulasan aplikasi sosial media Tiktok, X/Twiter dan Instagram yang ada pada Play Store dikumpulkan melalui proses *web scraping*. Teknik ini menggunakan *source code* untuk proses scraping menggunakan *Google colab*, dari proses tersebut akan mendapatkan data sentimen review komentar dalam format csv. Pada tahap ini data yang dikumpulkan sebanyak 15.600 dari bulan januari sampai juni data ulasan untuk masing-masing aplikasi social media. Data Ulasan *Social media* yang dikumpulkan berdasarkan aplikasi dan bulan. Ulasan positif dan negatif dapat dianalisis melalui diagram batang yang dimana diagram ini menampilkan jumlah bulanan label positif dan negatif pada aplikasi social media dari Januari hingga Juni 2024. Pada sumbu vertical (*y-axis*), jumlah label berkisar dari 0 hingga 2000, sementara sumbu horizontal (*x-axis*)

menunjukkan bulan dari Januari hingga Juni 2024. Setiap bulan diwakili oleh dua batang biru untuk label positif dan merah muda untuk label negatif. Berikut adalah diagram analisis ulasan positif serta negatif atas aplikasi *social media*.



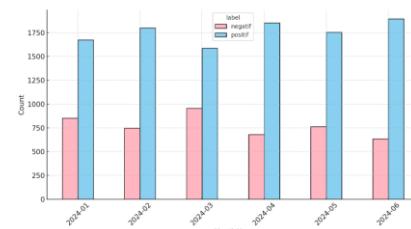
Gambar 2. Diagram Ulasan Positif dan Negatif Pada Aplikasi Tiktok

Berdasarkan gambar diagram diatas dapat dilihat bahwa dari bulan januari hingga juni jumlah ulasan positif lebih tinggi dibandingkan ulasan negatif dimana ulasan positif memiliki jumlah stabil antara 1.300 hingga 1.800 dan ulasan negatif sekitar 600 hingga 800. Dari diagram ini, terlihat bahwa ulasan positif lebih tinggi dibandingkan ulasan negatif untuk aplikasi Tiktok.



Gambar 3. Daigram Ulasan Positif dan Negatif Pada Aplikasi Tiktok

Pada diagram Twitter, terdapat pola yang berbeda dibandingkan Instagram. Ulasan negatif secara konsisten lebih tinggi daripada ulasan positif. Pada bulan Januari 2024, ulasan negatif mencapai angka tertinggi sebesar 1627, sementara ulasan positif berada di angka 875. Tren ini menunjukkan bahwa pengguna Twitter cenderung lebih kritis atau tidak puas dengan layanan atau produk di platform ini selama periode tersebut.



Gambar 4. Diagram Ulasan Positif dan Negatif pada Aplikasi Instagram

Diagram ini menunjukkan jumlah bulanan label positif dan negatif pada Instagram dari Januari hingga Juni 2024, dengan batang biru untuk label positif dan merah muda untuk label negatif. Label positif secara



konsisten mendominasi setiap bulan, dengan jumlah berkisar antara 1.500 hingga 1.800, sementara label negatif tetap stabil di sekitar 600 hingga 900.

B. Tahap Preprocessing

Setelah pengumpulan data, dilakukan preprocessing menggunakan Google Colab dan Python untuk membersihkan teks. Tahapannya meliputi *case folding*, mengubah huruf kapital menjadi kecil, *tokenizing* memecah teks menjadi kata-kata, *filtering* menghapus stopwords yang tidak signifikan seperti *yg*, *dg*, *tdk*, *ny*, serta *stemming* mengembalikan kata ke bentuk dasar dengan memangkas imbuhan. Proses ini bertujuan menyederhanakan teks agar siap digunakan dalam analisis.

Setelah melalui tahap Preprocessing jumlah kata mengalami pengurangan untuk menyingkirkan kata-kata yang mungkin kerap muncul tetapi tak berdampak signifikan pada dokumen. Dari 117.638 kata yang ada di data ulasan aplikasi Tiktok setelah dilakukan tahap Preprocessing memperoleh kata menjadi 114.259, kemudian 125.952 kata yang ada pada ulasan aplikasi Twitter setelah dilakukan tahap preprocessing memperoleh kata menjadi 122.323, dan untuk data ulasan aplikasi Instagram dari 109.418 kata setelah dilakukan tahap preprocessing memperoleh kata menjadi 106.929. Kata-kata tersebut dianggap penting dan memiliki pengaruh yang signifikan. Karena terjadi pengurangan kata dalam tahap preprocessing maka akan ada data ulasan yang berkurang juga. Dari 15.600 data ulasan aplikasi Tiktok yang ada berkurang menjadi 15.258, kemudian dari 15.600 data ulasan aplikasi Twitter yang ada berkurang menjadi 15.272, dan untuk aplikasi instagram dari 15.600 data ulasan berkurang menjadi 15.187 data ulasan.

C. Tahap Pembagian Data

Sesudah tahap preprocessing selesai, langkah berikutnya yakni data dijadikan data training serta data testing. Data sesudah diproses memakai tahapan preprocessing akan diklasifikasikan menjadi data training serta data testing menggunakan *Python*. Pada tahap ini data training dibagi sebesar 70% sedangkan data testing dibagi menjadi 30%.

D. Hasil Klasifikasi SVM dan Decision Tree

Setelah preprocessing, tahap selanjutnya adalah klasifikasi dengan menggunakan algoritma *Support Vector Machine (SVM)* dan *Decision Tree*. Pengujian dilakukan dengan nilai fold 2 hingga 10 untuk mencari performa terbaik. SVM diuji dengan tiga jenis kernel, yaitu *linear*, *radial*, dan *polynomial*, yang berfungsi meminimalkan error, sedangkan Decision Tree memisahkan data ke dalam subset berdasarkan fitur untuk membentuk pohon keputusan. Hasil uji coba pada tiga aplikasi media sosial menunjukkan performa masing-masing kernel *SVM* dibandingkan Decision Tree.

Tabel 2 Hasil Klasifikasi SVM Pada Aplikasi Tiktok

Nilai K-Folds	Linear	Radial	Polynomial
2	83.02%	83.71%	65.85%
3	83.56%	83.72%	67.35%

Vol.16 no.2 | Desember 2025

EXPLORE : ISSN: 2087-2062, Online ISSN: 2686-181X / DOI: <http://dx.doi.org/10.36448/jsit.v16i2.4455>



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

4	83.99%	84.05%	67.95%
5	83.99%	84.32%	68.45%
6	83.99%	84.33%	68.22%
7	83.83%	84.40%	68.71%
8	83.84%	84.51%	68.75%
9	83.72%	84.61%	68.74%
10	83.72%	84.51%	68.99%

Tabel 3 Hasil Klasifikasi SVM Pada Aplikasi X/Twitter

Nilai K-Folds	Linear	Radial	Polynomial
2	85.31%	85.85%	64.87%
3	85.57%	86.54%	66.38%
4	85.57%	86.66%	67.17%
5	85.72%	86.66%	67.62%
6	85.76%	86.23%	67.67%
7	85.66%	86.81%	67.93%
8	85.71%	86.45%	68.03%
9	85.83%	86.73%	68.12%
10	85.83%	86.82%	68.24%

Tabel 4 Hasil Klasifikasi SVM Pada Aplikasi Instagram

Nilai K-Folds	Linear	Radial	Polynomial
2	85.00%	85.17%	76.76%
3	85.66%	85.43%	76.64%
4	85.47%	85.53%	77.65%
5	85.41%	85.73%	77.65%
6	85.47%	85.72%	77.99%
7	85.49%	85.71%	77.69%
8	85.30%	85.63%	77.96%
9	85.34%	85.71%	78.13%
10	85.23%	85.71%	78.18%

Pada ketiga tabel di atas ditampilkan hasil eksperimen dari tiga aplikasi media sosial yang menggunakan metode *SVM*. Nilai akurasi tertinggi untuk aplikasi TikTok adalah 84.61%, dicapai dengan kernel tipe radial pada *k-fold* 9. Untuk aplikasi Twitter, nilai akurasi tertinggi adalah 86.82%, juga dengan kernel tipe radial pada *k-fold* 10. Sedangkan untuk aplikasi Instagram, nilai akurasi tertinggi adalah 85.73%, dicapai dengan kernel tipe *radial* pada *k-fold* 5.

Hasil dari uji coba tiga aplikasi *social media* menggunakan metode *Decision Tree* yang mana dijalankan melalui *k-fold* 2 hingga 10 menghasilkan nilai berikut:

Tabel 5 Hasil Klasifikasi Decision Tree Pada Aplikasi Tiktok

Nilai K-Folds	Hasil
2	77.51%
3	77.98%
4	77.63%

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

5	77.98%
6	77.81%
7	77.84%
8	77.90%
9	77.78%
10	77.58%

Tabel 6 Hasil Klasifikasi Decision Tree Pada Aplikasi X/Twitter

Nilai K-Folds	Hasil
2	79.47%
3	79.51%
4	79.75%
5	79.90%
6	79.82%
7	79.83%
8	80.07%
9	80.14%
10	79.98%

Tabel 7 Hasil Klasifikasi Decision Tree Pada Instagram

Nilai K-Folds	Hasil
2	81.26%
3	80.75%
4	81.63%
5	81.61%
6	81.39%
7	81.47%
8	81.33%
9	81.58%
10	81.71%

TikTok memiliki akurasi terbaik pada *k-fold* 3 dan 5 dengan nilai sebesar 77.98%. Tabel kedua, yang menampilkan hasil dari aplikasi Twitter, menunjukkan akurasi terbaik pada *k-fold* 9 dengan nilai sebesar 80.14%. Tabel ketiga, yang menunjukkan hasil dari aplikasi Instagram, memiliki akurasi terbaik pada *k-fold* 10 dengan nilai sebesar 81.71%.

E. Evaluasi

Berdasarkan hasil klasifikasi yang telah dilakukan maka menghasilkan model dengan akurasi tertinggi. Pengujian dilakukan dengan *operator cross-validation* menggunakan *k-fold* dari 2 hingga 10. Dalam eksperimen dengan metode Support Vector Machine dan Decision Tree.

Perbandingan hasil eksperimen menegaskan bahwasanya nilai akurasi paling baik pada uji *SVM* dan *Decision Tree* nampak di table berikut:

Tabel 8 Tabel Perbandingan Algoritma Pada Aplikasi Tiktok

Metrik	SVM (Kernel: rbf/radial)	Decision Tree
Akurasi	84.61%	77.98%
Precision (Negatif)	0.71	0.60
Precision (Positif)	0.87	0.84
Recall (Negatif)	0.68	0.57
Recall (Positif)	0.89	0.86
F1-Score (Negatif)	0.69	0.59
F1-Score (Positif)	0.88	0.85
Support (Negatif)	1312	1265
Support (Positif)	3266	3313
Macro Avg (Precision)	0.79	0.72
Macro Avg (Recall)	0.78	0.71
Macro Avg (F1-Score)	0.79	0.72
Weighted Avg (Precision)	0.83	0.77
Weighted Avg (Recall)	0.83	0.78
Weighted Avg (F1-Score)	0.83	0.78
Matriks Kebingungan	[[889, 476], [314, 2952]]	[[722, 543], [472, 2841]]

Berdasarkan hasil perbandingan, *SVM* dengan *kernel RBF* terbukti lebih unggul dibandingkan Decision Tree pada klasifikasi ulasan aplikasi Tiktok. *SVM* mencapai akurasi 84,61% dengan precision, recall, dan *F1-score* yang lebih tinggi di kedua kelas dibandingkan Decision Tree yang hanya memperoleh akurasi 77,98%. Rata-rata makro, rata-rata berbobot, serta matriks kebingungan juga menunjukkan bahwa *SVM* menghasilkan prediksi yang lebih akurat, sehingga dapat disimpulkan bahwa *SVM* dengan *kernel RBF* memberikan performa terbaik.

Tabel 9 Tabel Perbandingan Algoritma Pada Aplikasi X/Twitter

Metrik	SVM (Kernel: rbf/radial)	Decision Tree
Akurasi	86.82%	80.14%
Precision (Negatif)	0.87	0.80
Precision (Positif)	0.87	0.81
Recall (Negatif)	0.91	0.81
Recall (Positif)	0.82	0.80
F1-Score (Negatif)	0.89	0.81



F1-Score	0.84	0.81	Weighted Avg	0.85	0.82
(Positif)			(F1-Score)		
Support	2627	2281	Matriks	[[1017, 351],	[[663, 431],
(Negatif)			Kebingungan	[330, 2859]]	[374, 3089]]
Support (Positif)	1955	2315			
Macro Avg (Precision)	0.87	0.81			
Macro Avg (Recall)	0.87	0.81			
Macro Avg (F1-Score)	0.80	0.81			
Weighted Avg (Precision)	0.87	0.81			
Weighted Avg (Recall)	0.87	0.81			
Weighted Avg (F1-Score)	0.87	0.81			
Matriks	[[2313, 314],	[[1858, 423],			
Kebingungan	[317, 1638]]	[464, 1851]]			

Hasil perbandingan menunjukkan bahwa *SVM* dengan kernel RBF lebih unggul dibandingkan *Decision Tree* pada klasifikasi ulasan Twitter. *SVM* mencapai akurasi 86,82% dengan *precision*, *recall*, dan *F1-score* yang lebih tinggi pada kedua kelas, serta rata-rata makro dan berbobot yang juga lebih baik. Matriks kebingungan memperlihatkan *SVM* menghasilkan prediksi lebih akurat dengan kesalahan lebih sedikit, sehingga dapat disimpulkan bahwa *SVM* dengan kernel RBF memberikan performa terbaik.

Tabel 10 Tabel Perbandingan Algoritma Pada Aplikasi Instagram

Metrik	SVM (Kernel: rbf/radial)	Decision Tree
Akurasi	85.73%	81.71%
Precision (Negatif)	0.76	0.64
Precision (Positif)	0.89	0.88
Recall (Negatif)	0.74	0.61
Recall (Positif)	0.90	0.89
F1-Score (Negatif)	0.75	0.62
F1-Score (Positif)	0.89	0.88
Support (Negatif)	1368	1094
Support (Positif)	3189	3463
Macro Avg (Precision)	0.82	0.76
Macro Avg (Recall)	0.82	0.75
Macro Avg (F1-Score)	0.82	0.75
Weighted Avg (Precision)	0.85	0.82
Weighted Avg (Recall)	0.85	0.82

Berdasarkan hasil perbandingan, *SVM* dengan kernel RBF terbukti lebih unggul dibandingkan *Decision Tree* pada klasifikasi ulasan aplikasi Instagram. *SVM* mencapai akurasi 85,73% dengan *precision*, *recall*, dan *F1-score* yang lebih tinggi di kedua kelas dibandingkan *Decision Tree* yang hanya memperoleh akurasi 81,71%. Rata-rata makro, rata-rata berbobot, serta matriks kebingungan juga menunjukkan bahwa *SVM* menghasilkan prediksi yang lebih akurat dengan kesalahan lebih sedikit, sehingga dapat disimpulkan bahwa *SVM* dengan kernel RBF memberikan performa terbaik.

4. Kesimpulan

Berdasarkan temuan analisis antara metode algoritma Support vector Machine dan Decision Tree pada data ulasan aplikasi social media, dapat disimpulkan bahwa:

1. Hasil analisis sentimen data ulasan pada aplikasi social media seperti Tiktok sebesar 11.070 untuk ulasan positif, 4.188 untuk ulasan negatif dan untuk aplikasi Twitter memiliki 7.589 ulasan positif dan 7.729 untuk ulasan negatif, sementara itu untuk aplikasi Instagram memiliki 11.546 untuk ulasan positif dan 3.641 untuk ulasan negatif
2. Hasil akurasi data ulasan pada aplikasi social media memanfaatkan metode Support Vector Machine sebesar 84.61% untuk aplikasi Tiktok, 86.82% Untuk aplikasi Twitter dan 85.73% untuk aplikasi Instagram, sedangkan metode Decision Tree sebesar 77.98% untuk aplikasi Tiktok, 80.14% untuk aplikasi Twitter, dan 81.71% untuk aplikasi Instagram. Hasil Akurasi, dari *SVM* secara konsisten menegaskan akurasi unggul dibandingkan Decision Tree di semua dataset aplikasi social media.
3. *SVM* umumnya memiliki nilai yang lebih tinggi untuk kedua kelas (negatif dan positif) dibandingkan Decision Tree. Ini menunjukkan bahwa *SVM* lebih baik dalam mengidentifikasi kedua kelas dengan lebih sedikit kesalahan. Precision untuk kelas positif pada *SVM* berkisar antara 0.87 hingga 0.89, sedangkan pada Decision Tree berkisar antara 0.81 hingga 0.88.
4. *SVM* memiliki lebih sedikit false positives dan false negatives dibandingkan Decision Tree. Misalnya, pada salah satu dataset, *SVM* memiliki matriks kebingungan [[889, 476], [314, 2952]], sementara Decision Tree memiliki [[772, 543], [472, 2841]]. Ini mengkonfirmasi bahwa *SVM* lebih unggul dalam meminimalkan kesalahan klasifikasi.
5. Konsistensi Performa, *SVM* menunjukkan performa yang lebih konsisten di berbagai dataset, dengan tetap mempertahankan akurasi dan metrik evaluasi yang tinggi, sementara performa Decision Tree lebih bervariasi. Secara keseluruhan, *SVM* adalah model yang lebih andal dan efektif untuk dataset yang diberikan, menegaskan performa optimal perihal



akurasi, precision, recall, F1-score, dan analisis matriks kebingungan dibandingkan Decision Tree.

5. Daftar Pustaka

- [1] D. Affandi and I. Wijayani, “Social Media as Self Existence in Students Using Tiktok Applications,” *Dawatuna J. Commun. Islam. Broadcast.*, vol. 2, no. 3, pp. 300–311, Aug. 2022, doi: 10.47467/dawatuna.v2i3.2108.
- [2] A. Aldila Safitri, A. Rahmadhany, and I. Irwansyah, “Penerapan Teori Penetrasi Sosial pada Media Sosial: Pengaruh Pengungkapan Jati Diri melalui TikTok terhadap Penilaian Sosial,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 3, no. 1, pp. 1–9, Jan. 2021, doi: 10.47233/jtekstis.v3i1.180.
- [3] F. A. Girnanfa and A. Susilo, “Studi Dramaturgi Pengelolaan Kesan Melalui Twitter Sebagai Sarana Eksistensi Diri Mahasiswa di Jakarta,” *J. New Media Commun.*, vol. 1, no. 1, pp. 58–73, Apr. 2022, doi: 10.55985/jnmc.v1i1.2.
- [4] N. Istiqomah and F. Novika, “Sentiment Analysis Penyedia layanan Asuransi dari Media Sosial Twitter,” vol. 18, no. 1.
- [5] W. Surya Gemilang, P. Purwantoro, and C. Carudin, “Analisis Sentimen Pengguna Instagram Pada Calon Presiden 2024 Menggunakan Algoritma Support Vector Machine,” *JATI J. Mbs. Tek. Inform.*, vol. 7, no. 4, pp. 2849–2855, Jan. 2024, doi: 10.36040/jati.v7i4.7256.
- [6] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiasari, “Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE,” *AITI*, vol. 18, no. 2, pp. 173–184, Nov. 2021, doi: 10.24246/aiti.v18i2.173-184.
- [7] V. K. S. Que, A. Iriani, and H. D. Purnomo, “Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization,” *J. Nas. Tek. Elektro Dan Teknol. Inf.*, vol. 9, no. 2, pp. 162–170, May 2020, doi: 10.22146/jnteti.v9i2.102.
- [8] M. Alfarizi, M. Rizqy, R. I. Ghufroni, R. D. Saputra, and F. Kurniawan, “Analisis Sentimen Persepsi Publik Terhadap Kasus Bullying Siswa Cilacap Menggunakan Pendekatan Machine Learning,” vol. 4, no. 3, 2023.
- [9] C. F. Hasri and D. Alita, “Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter,” vol. 3, no. 2.
- [10] A. Mustofa and R. Novita, “Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter,” *Build. Inform. Technol. Sci. BITS*, vol. 4, no. 1, Jun. 2022, doi: 10.47065/bits.v4i1.1628.
- [11] F. Pratama, M. Nasir, and S. Sauda, “Implementasi Metode Klasifikasi Dengan Algoritma Support Vector Machine Untuk Menentukan Stok Persediaan Barang Pada Koperasi Karyawan Pangan Utama,” *J. Softw. Eng. Ampera*, vol. 1, no. 2, pp. 71–81, May 2020, doi: 10.51519/journalsea.v1i2.46.
- [12] M. H. Wicaksono, M. D. Purbolaksono, and S. A. Faraby, “Perbandingan Algoritma Machine Learning untuk Analisis Sentimen Berbasis Aspek pada”.
- [13] I. Arfyanti, M. Fahmi, and P. Adytia, “Penerapan Algoritma Decision Tree Untuk Penentuan Pola Penerima Beasiswa KIP Kuliah,” *Build. Inform. Technol. Sci. BITS*, vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2275.
- [14] [14] R. Ridho, “Klasifikasi Diagnosis Penyakit Covid-19 Menggunakan Metode Decision Tree”.
- [15] [15] R. A. Saputra, D. P. Ray, and F. Irwiensyah, “Analisis Sentimen Aplikasi Tokocrypto Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Naïve Bayes”.

