

ANALISIS SENTIMEN PENGGUNA TIK TOK SHOP MENGGUNAKAN ALGORITMA SVM

Qois Al Qorni^{1*}, Tata Sutabri²⁾

^{1,2} Magister Teknik Informatika, Universitas Bina Darma
email: ¹ qoisalqorni501@gmail.com, ² tatasutabri@gmail.com

Abstract

In today's digital landscape, social media platforms like TikTok Shop have gained immense popularity as online shopping hubs. This study advocates the utilization of Support Vector Machine (SVM) methodology to assess customer sentiment concerning products available on TikTok Shop. The primary objective is to assess SVM's efficacy in categorizing customer sentiments (positive, negative) based on their responses to product reviews. The research findings revealed a classification accuracy of 78.66%, precision of 78.01%, recall of 99.09%, and an F1 score of 87.30%. While these results showcase a commendable performance, the study recognizes room for improvement. Consequently, further enhancements are recommended, particularly in refining preprocessing techniques for both data and modeling. These additional preprocessing steps are anticipated to significantly enhance sentiment classification accuracy, ensuring more reliable outcomes and providing deeper insights into customer preferences within the realm of online shopping.

Keywords: Analisis, Sentimen, SVM, TikTok, Shop

1. PENDAHULUAN

Di era komputerisasi yang berkembang pesat ini, hiburan online telah menjadi bagian penting dari rutinitas rutin banyak orang, membentuk cara kita bekerja sama, berbagi data, dan berbelanja. Dorongan dalam inovasi data melibatkan kelompok usia yang berbeda-beda, termasuk anak muda, remaja, orang dewasa, dan, yang mengejutkan, orang tua yang tetap terdorong untuk mengikuti dan mengetahui kemajuan dalam inovasi data. Hal ini terutama terlihat di bidang hiburan virtual, bisnis online, dan kemajuan komputer lainnya yang sepenuhnya terhubung melalui web [1]. Salah satu media informasi yang banyak digunakan adalah Tik Tok.

Sistem Informasi merupakan sebuah sistem yang membantu menghasilkan informasi yang dibutuhkan dengan cara yang terbaik. Sistem ini mengelola data dengan efisien dan menghasilkan informasi yang berkualitas. Ini berarti membantu organisasi untuk melakukan kegiatan atau proses mereka dengan lebih baik dan efisien [2], [3].

Aplikasi TikTok adalah platform hiburan berbasis web terbaru yang memungkinkan klien untuk membuat perkamen dan menawarkan rekaman menarik, serta berkolaborasi melalui segmen komentar dan pertemuan rahasia. Aplikasi ini menawarkan berbagai hiasan menarik dan mudah digunakan, memungkinkan siapa pun membuat

rekaman kreatif. Variabel ini membuat Tik-Tok menjadi sangat populer, dengan banyak pengguna yang menggunakannya. Hal ini ditunjukkan dengan tingginya penilaian yang diperoleh aplikasi ini di toko aplikasi Play Store, mencapai 4,6 dari 5 bintang, dan mendapatkan dukungan dari sekitar 27.827 klien di seluruh dunia. Pemeriksaan dengan aplikasi serupa, Musicaly, menunjukkan perbedaan ketenaran yang luar biasa, dengan peringkat 3,5 dari 5 bintang, skenario kasus terbaik, dan hanya 4.100 klien. Tik Tok pastinya akan memikat klien dengan beragam elemen imajinatifnya, menjadikannya salah satu aplikasi utama di ranah hiburan berbasis web saat ini. [4].

Dari banyaknya inovasi teknologi informasi di bidang media sosial, Tik Tok merupakan salah satu aplikasi media sosial yang sekaligus merangkap menjadi e-commerce (Tik Tok Shop). Dalam konteks TikTok Shop, Sistem Informasi membantu mengelola produk, memproses pembayaran, dan memberikan informasi relevan kepada pengguna. Dengan cara ini, TikTok Shop dapat membantu pengguna menemukan produk yang mereka inginkan dan membantu pedagang mengelola inventaris dan penjualan mereka dengan lebih baik. Sistem Informasi di TikTok Shop juga dapat memberikan analisis data untuk

memahami preferensi pelanggan, memperbaiki pengalaman pengguna, dan mengoptimalkan proses bisnis secara keseluruhan. Dari fungsional dan efektifitas yang sangat banyak itu mengakibatkan berbagai opini dan komentar miring dari masyarakat sekitar. Semenjak tingginya minat masyarakat berbelanja online, toko offline terutama pasar merasakan dampak sepi pembeli.

Analisis sentimen adalah cabang studi yang mengkaji pendapat, sikap, penilaian, dan emosi individu terhadap topik, layanan, produk, organisasi, atau kegiatan tertentu [5]. Analisis sentimen bertujuan untuk mengidentifikasi dan mengklasifikasikan konten teks dalam dataset sebagai positif, negatif, atau netral [6]. Proses analisis sentimen bertujuan untuk mengevaluasi apakah opini atau komentar terhadap suatu isu cenderung positif atau negatif. Hasil analisis sentimen ini dapat digunakan sebagai panduan untuk meningkatkan layanan atau memperbaiki kualitas produk, serta memberikan wawasan yang berharga terhadap persepsi publik terhadap suatu entitas atau topik tertentu.

Untuk membantu sistem pengelompokan informasi survei penilaian pada aplikasi Tik Tok Shop, diperlukan strategi AI yang dapat menginteraksikan informasi tersebut. Support Vector Machine (SVM) adalah kerangka kerja AI yang melibatkan spekulasi sebagai elemen lurus dalam kemampuan berlapis dan disiapkan menggunakan perhitungan pembelajaran sehubungan dengan perampingan hipotesis. [7]. Support Vector Machine (SVM) adalah salah satu dari sekian banyak teknik yang sering diterapkan untuk berbagai jenis pemeriksaan di bidang penanganan informasi dan penambangan teks, karena mampu menunjukkan kinerja yang efektif dan lebih baik. [8].

Maka pada pengujian kali ini strategi tersebut akan dijalankan dan diikuti hasil ketepatannya dengan menggunakan model Support Vector Machine (SVM) menggunakan peralatan Google Colab dengan bahasa pemrograman Python dan hasil karakterisasinya dapat membantu menentukan dan mengatur jenis perasaan dengan menggunakan perhitungan AI.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis merupakan tindakan belajar mencari contoh atau strategi anggapan yang dihubungkan dengan pengujian logis terhadap informasi untuk menentukan klasifikasi, hubungan antar bagian, dan

mengikat keseluruhan. [9].

Analisis sentimen ialah cara paling umum untuk menentukan perspektif atau perasaan komentator dan mengumpulkan kata-kata ekstrem dalam laporan atau kalimat, sehingga kelas tidak sepenuhnya menganggap opini baik atau pesimistis. Saat ini, para ilmuwan secara luas menggunakan pemeriksaan opini sebagai bagian dari eksplorasi dalam rekayasa perangkat lunak. Organisasi hiburan online, seperti Twitter, sebagian besar digunakan dalam menguji perasaan untuk menentukan kebijaksanaan publik [10].

2.2 Text Mining

Text Mining merupakan interaksi untuk memisahkan data unggulan dari pesan. Data yang baik seringkali diperoleh dengan memperhatikan contoh dan pola melalui penyelidikan faktual. Pada proses text mining terdapat langkah pembobotan kata yang diharapkan dapat memberikan nilai atau bobot pada kata-kata yang ada pada suatu record. Bobot yang diberikan pada kata-kata ini bergantung pada teknik yang digunakan. Dalam proses pembobotan kata terdapat perhitungan yang berbeda-beda seperti TF, IDF, RF, TF-IDF, TF.RF, dan WIDF. Penghitungan ini berperan penting dalam menentukan relevansi dan makna kata dalam kumpulan data teks [11].

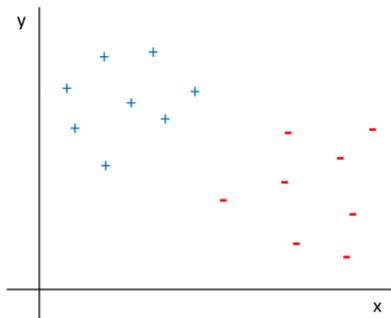
2.3 Algoritma Machine Learning

Machine Learning ialah salah satu cabang dari ilmu kecerdasan buatan, khususnya yang menekuni tentang gimana komputer sanggup belajar dari informasi buat tingkatkan kecerdasannya. Machine learning mempunyai fokus pada pengembangan suatu sistem yang sanggup belajar sendiri buat memutuskan suatu, tanpa wajib kesekian kali deprogram oleh manusia. Dengan begitu mesin tidak Cuma dapat menciptakan ketentuan buat perilaku maksimal dalam pengambilan keputusan, tetapi pula dapat menyesuaikan diri dengan pergantian yang terjalin [12].

Dalam eksplorasi ini digunakan perhitungan karakterisasi Support Vector Machine (SVM). SVM adalah perhitungan AI terarah yang dapat digunakan untuk kesulitan karakterisasi dan relaps. Meskipun demikian, ini terutama digunakan dalam masalah pengaturan. Dalam perhitungan SVM, kita dapat mengamati setiap data yang diinginkan sebagai sebuah titik dalam ruang berlapis n (di mana n adalah jumlah sorotan eksplisit yang kita miliki), di mana nilai setiap komponen merupakan nilai arah tertentu. Kemudian, pada saat itu, kami melakukan pengurutan dengan mencari hyperplane yang mampu mengisolasi kedua kelas dengan baik.

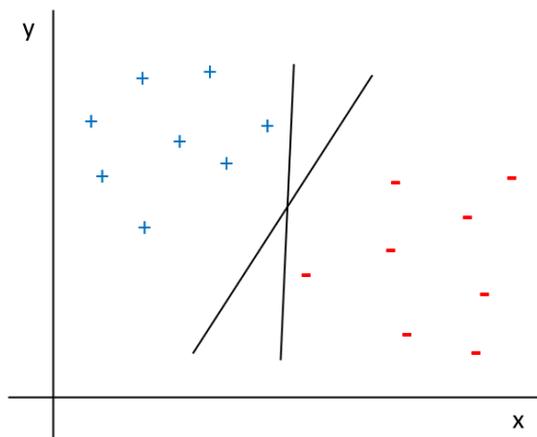
Cara kerja SVM, Pengoperasian algoritma SVM dapat dipahami dengan sebuah contoh.

Misalnya, ada Kumpulan data dengan dua label, positif dan negative, Kumpulan data tersebut memiliki dua fitur x_1 dan x_2 . Kita akan melakukan proses klasifikasi terhadap sepasang koordinat x_1 sebagai positif dan x_2 sebagai negative. Dapat dilihat pada Gambar 1.



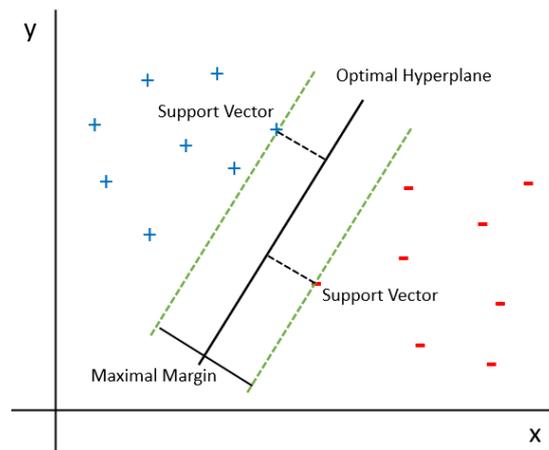
Gambar 1. Ruang 2D

Gambar 1 merupakan ruang 2D, dapat dengan mudah memisahkan data menggunakan garis. Dapat dilihat pada Gambar 2.



Gambar 2. Garis Pemisah

Gambar diatas merupakan contoh pemanfaatan garis untuk memisahkan 2 data. Akan tetapi, data yang dipisahkan tidak optimal atau tidak terbagi dengan merata. Algoritma SVM berperan penting dalam mencari garis keputusan terbaik (Hyperplane) dengan mencari titik terdekat dari garis kedua data. Titik-titik ini disebut vektor dukungan. Jarak antara vektor dan hyperplane disebut margin. Tujuan SVM adalah memaksimalkan margin tersebut. Hyperplane dengan tepi maksimum disebut hyperplane optimal. Dapat dilihat pada Gambar 3.



Gambar 3. Hyperplane SVM

2.4 Data Mining

Data merupakan sekumpulan fakta atau informasi yang dapat diukur, dihitung, atau diolah. Data dapat berupa angka, kata-kata, gambar, atau bentuk lainnya. Dalam konteks komputasi, data adalah representasi digital dari informasi, yang dapat disimpan, diolah, dan ditransmisikan oleh computer.

Sedangkan data mining, merupakan proses yang mencakup pengumpulan dan penggunaan data historis untuk mengidentifikasi pola, keteraturan, atau hubungan dalam dataset besar. Hasil dari data mining ini dapat digunakan untuk meningkatkan pengambilan keputusan di masa mendatang. Saat ini, istilah pengenalan pola (pattern recognition) jarang digunakan karena konsep ini sudah termasuk dalam ranah data mining [13].

2.5 Confusion Matrix

Confusion Matrix adalah denah meja yang menggunakan empat perpaduan berbeda antara kualitas yang diantisipasi dan asli. Penilaian pelaksanaan penanganan ini memanfaatkan kemampuan F1 Score dan ketepatan. Skor F1 adalah kemampuan yang melihat tinjauan dan keakuratan, sedangkan ketepatan menggambarkan seberapa tepat model saat ini mengkarakterisasi informasi secara akurat. Kisi kekacauan adalah strategi yang biasanya digunakan untuk melakukan cara paling umum dalam menghitung ketepatan dalam ide penambangan informasi [14]. Contoh desain Confusion Matrix terdapat pada Tabel 1.

| Predict Class | Actual Class | |
|---------------|---------------|---------------|
| | Positif | Negatif |
| Positif | True Positif | False Negatif |
| Negatif | False Negatif | True Negatif |

Diperoleh empat nilai keluaran berdasarkan Tabel 1:

Recall

Fungsi yang digunakan untuk Perhitungan dalam mengetahui berapa besar persentase studi kasus data yang benar diidentifikasi sebagai benar.

$$Recall = TP / TP+FN \quad (1)$$

Accuracy

Hasil akhir akurasi prediksi yang diklasifikasi menggunakan data lainnya untuk mengetahui keakuratan permodelan yang digunakan.

$$Accuracy = TP+TN / TP+FP+FN+TN \quad (2)$$

Precision

Fungsi perbandingan diantara hasil akhir klasifikasi TP di klasifikasi sebagai nilai positif dengan semua prediksi nilai positif pula.

$$Precision = TP / TP+FP \quad (3)$$

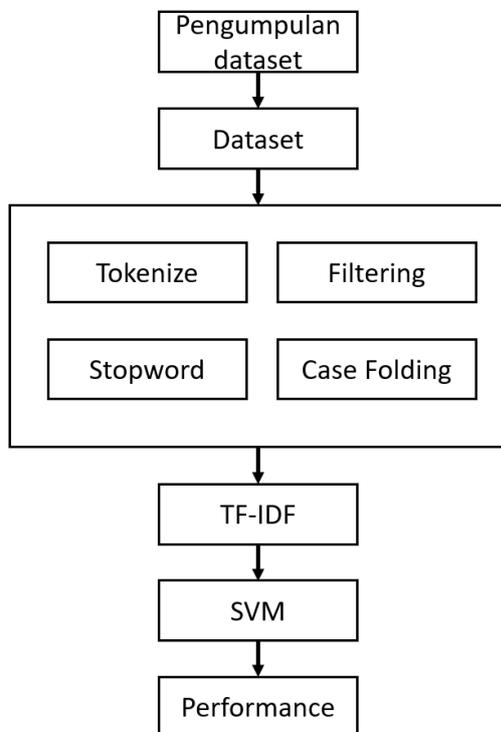
F1 Score

Fungsi perbandingan perhitungan rata-rata precision dan recall yang dibobotkan

$$F1 \text{ Score} = 2 * recall * precision / recall + precision \quad (4)$$

3. METODE PENELITIAN

Penelitian ini menggunakan beberapa tahap proses mulai dari pengambilan data, pemrosesan data hingga test performance menggunakan algoritma SVM, terlihat pada gambar berikut:



Gambar 4. Alur Penelitian

Langkah pertama yang dilakukan adalah

pengumpulan data. Pada penelitian ini data berasal dari komentar twitter terhadap tutupnya Tik Tok Shop, data yang dicari menggunakan kata kunci “Tik Tok Shop Tutup” yang dilakukan pada 1 Oktober 2023 dan data yang didapatkan sebanyak 500 data. Pengambilan data dilakukan dengan memanfaatkan API Token Twitter, kemudian data di crawling menggunakan Tools Google Colab, setelah proses crawling dataset selesai, selanjutnya proses penyimpanan data dengan format CSV.

Selanjutnya, dataset yang telah didapatkan di olah kembali dengan hanya memanfaatkan kolom text atau komentar saja. Setelah memiliki data dengan kolom text, ditambahkan lagi kolom baru dengan nama label untuk memberi sentiment positif atau negative dari komentar yang ada.

Tahap ketiga Preprocessing Data, suatu proses cleaning data. Adapun tahapan dari preprocessing data adalah Tokenize, proses untuk memecah karakter atau frasa menjadi potongan-potongan kata yang terpisah. Selanjutnya Case Folding, proses merubah kata atau kalimat yang memiliki huruf besar akan disama ratakan menjadi huruf kecil semua. Kemudian Filtering, proses membuang beberapa kata yang tidak diperlukan setelah proses tokenize. Dan terakhir Stemming, proses membuang kata tambahan seperti dan, atau, namun, akan, tetapi, dll. Karena kata tersebut tidak memiliki makna khusus.

Tahap selanjutnya TF-IDF, melakukan proses konfersi perhitungan kemunculan kata. TF-IDF berguna menentukan frekuensi sebuah kata dalam dokumen dan memberikan bobot, setelah kata diberi bobot, kata yang memiliki bobot tertinggi dikatakan sebagai kata yang sering muncul.

Terakhir menggunakan SVM sebagai algoritma classification untuk membuat model dan melihat hasil performa akurasi klasifikasi.

- 3223-4_13.
- [6] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," *Expert Syst Appl*, vol. 40, no. 10, pp. 4065–4074, Aug. 2013, doi: 10.1016/j.eswa.2013.01.001.
- [7] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, "SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM," *Jurnal Data Mining dan Sistem Informasi*, vol. 2, no. 1, 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [8] D. Alita, Y. Fernando, and H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," *Jurnal Tekno Kompak*, vol. 14, no. 2, 2020, doi: 10.33365/jtk.v14i2.792.
- [9] G. Buana Putri, A. Khoir Diaz, and T. Sutabri, "Algoritma Machine Learning Untuk Klasifikasi Review Film ada IMDB 25 ALGORITMA MACHINE LEARNING UNTUK KLASIFIKASI REVIEW FILM PADA IMDB," *Dinamika Informatika*, vol. 15, no. 1, 2023.
- [10] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 1, 2021, doi: 10.30865/mib.v5i1.2580.
- [11] A. Deolika and E. Taufiq Luthfi, "ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING," *Jurnal Teknologi Informasi*, vol. 3, no. 2, 2019.
- [12] Putri Ramayanti and Tata Sutabri, "PERBANDINGAN ALGORITMA NAÏVE BAYES DAN SVM UNTUK ANALISIS PENYALAHGUNAAN KEJAHATAN CARDING," *Jurnal Informatika Teknologi dan Sains*, vol. 5, no. 1, 2023, doi: 10.51401/jinteks.v5i1.2231.
- [13] D. Novianti, "Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner," *Paradigma - Jurnal Komputer dan Informatika*, vol. 21, no. 1, 2019, doi: 10.31294/p.v21i1.4979.
- [14] M. F. Rahman, M. Ilham Darmawidjadja, and D. Alamsah, "KLASIFIKASI UNTUK DIAGNOSA DIABETES MENGGUNAKAN METODE BAYESIAN REGULARIZATION NEURAL NETWORK (RBNN)," 2017.