

Klasifikasi Minat Belanja Online pada TikTok Berdasarkan Konten Media Sosial Menggunakan Metode Naive Bayes

Puspita Wanny^{1,*}, Muhammad Iqbal²

Program Pasca Sarjana, Program Studi Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: ^{1,*}puspitawanny142@gmail.com, ²muhammadiqbalpb@gmail.com

Email Penulis Korespondensi: puspitawanny142@gmail.com

Abstrak—Perkembangan media sosial, khususnya TikTok, telah mendorong perubahan perilaku konsumen dalam belanja online melalui konten digital dan fitur TikTok Shop. Penelitian ini bertujuan untuk mengklasifikasikan minat belanja online pengguna TikTok berdasarkan jenis konten media sosial menggunakan metode Naive Bayes. Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining terhadap 100 data yang terdiri dari 50 data (50%) kategori Tertarik dan 50 data (50%) kategori Tidak Tertarik. Jenis konten yang dianalisis meliputi Promosi, Review Produk, Hiburan, Edukasi, dan Vlog. Hasil implementasi Naive Bayes menunjukkan nilai probabilitas prior $P(\text{Tertarik}) = 0,5$ dan $P(\text{Tidak Tertarik}) = 0,5$. Perhitungan probabilitas kondisional menghasilkan nilai posterior sebesar 0,25 untuk konten Promosi dan Review Produk sehingga diklasifikasikan sebagai Tertarik. Sebaliknya, konten Hiburan dan Edukasi memiliki nilai posterior sebesar 0,19, sedangkan konten Vlog sebesar 0,12, sehingga diklasifikasikan sebagai Tidak Tertarik. Hasil akhir menunjukkan bahwa 2 dari 5 jenis konten (40%) berpengaruh terhadap minat belanja, sedangkan 3 jenis konten (60%) tidak berpengaruh. Dengan demikian, metode Naive Bayes efektif digunakan untuk mengklasifikasikan minat belanja online berdasarkan jenis konten TikTok.

Kata kunci: TikTok; Minat Belanja Online; Konten Media Sosial; Data Mining; Naive Bayes

Abstract—The development of social media, especially TikTok, has driven changes in consumer behavior in online shopping through digital content and the TikTok Shop feature. This study aims to classify TikTok users' online shopping interests based on the type of social media content using the Naive Bayes method. This study uses a quantitative approach with data mining methods on 100 data consisting of 50 data (50%) in the Interested category and 50 data (50%) in the Not Interested category. The types of content analyzed include Promotion, Product Review, Entertainment, Education, and Vlog. The results of the Naive Bayes implementation show a prior probability value of $P(\text{Interested}) = 0.5$ and $P(\text{Not Interested}) = 0.5$. The conditional probability calculation produces a posterior value of 0.25 for Promotion and Product Review content, so they are classified as Interested. Conversely, Entertainment and Education content has a posterior value of 0.19, while Vlog content is 0.12, so it is classified as Not Interested. The final results showed that 2 of 5 content types (40%) influenced shopping interest, while 3 content types (60%) had no effect. Thus, the Naive Bayes method is effective for classifying online shopping interest based on TikTok content type.

Keywords: TikTok; Online Shopping Interest; Social Media Content; Data Mining; Naive Bayes

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi di era digital telah membawa perubahan fundamental dalam berbagai aspek kehidupan manusia, termasuk dalam bidang ekonomi dan perdagangan. Internet tidak hanya berfungsi sebagai sarana pertukaran informasi, tetapi juga telah menjadi medium utama dalam aktivitas transaksi ekonomi. Perubahan ini mendorong pergeseran pola belanja masyarakat dari sistem konvensional menuju sistem digital atau belanja online. Konsumen kini lebih memilih berbelanja secara online karena menawarkan kemudahan akses, efisiensi waktu, serta beragam pilihan produk dan layanan yang dapat diperoleh tanpa batasan ruang dan waktu [1].

Seiring dengan berkembangnya belanja online, media sosial memegang peranan yang sangat penting dalam membentuk perilaku konsumen. Media sosial tidak lagi hanya digunakan sebagai sarana komunikasi dan hiburan, tetapi juga telah berkembang menjadi media pemasaran digital yang strategis [2]. Platform media sosial memungkinkan pelaku usaha untuk mempromosikan produk secara langsung kepada konsumen melalui konten digital yang menarik, kreatif, dan interaktif. Konten media sosial yang efektif mampu memengaruhi persepsi, sikap, serta minat konsumen terhadap suatu produk atau jasa.

Salah satu platform media sosial yang mengalami pertumbuhan sangat pesat dalam beberapa tahun terakhir adalah TikTok. TikTok merupakan platform berbasis video pendek yang menawarkan pengalaman visual dan audio yang menarik bagi penggunaannya. Keunggulan TikTok terletak pada algoritma rekomendasi konten yang mampu menyesuaikan video dengan preferensi pengguna berdasarkan riwayat interaksi, sehingga konten promosi dapat menjangkau target audiens secara lebih tepat. Popularitas TikTok yang tinggi menjadikannya sebagai salah satu media yang potensial dalam memengaruhi perilaku konsumsi masyarakat, khususnya generasi muda. Perkembangan TikTok tidak hanya berdampak pada aspek hiburan, tetapi juga pada sektor perdagangan digital. Kehadiran fitur TikTok Shop memungkinkan pengguna untuk melakukan aktivitas belanja online secara langsung melalui aplikasi TikTok. Pengguna dapat melihat video promosi produk, membaca ulasan atau komentar dari pengguna lain, serta melakukan transaksi pembelian tanpa harus berpindah ke platform lain [3]. Integrasi antara konten hiburan dan fitur belanja ini menjadikan TikTok sebagai platform e-commerce berbasis media sosial yang unik dan memiliki daya tarik tersendiri.

Dalam konteks belanja online, minat belanja merupakan faktor penting yang menentukan apakah konsumen akan melakukan pembelian atau tidak. Minat belanja online dapat diartikan sebagai kecenderungan psikologis individu untuk melakukan pembelian produk atau jasa melalui media digital. Minat ini terbentuk melalui proses evaluasi informasi yang diterima konsumen, baik yang bersifat kognitif maupun afektif. Pada platform TikTok, informasi tersebut disajikan dalam

bentuk konten media sosial yang mencakup video promosi, caption, penggunaan musik, serta interaksi sosial berupa komentar, like, dan share[4].

Karakteristik konten media sosial memiliki peran yang signifikan dalam memengaruhi minat belanja online. Konten dengan kualitas visual yang baik, penyampaian pesan yang jelas, serta didukung oleh kreator atau influencer yang kredibel cenderung lebih mampu menarik perhatian konsumen. Pendekatan data mining menjadi salah satu metode yang relevan untuk menganalisis data perilaku konsumen dalam jumlah besar. Data mining memungkinkan peneliti untuk menggali pola, hubungan, dan informasi penting dari kumpulan data yang kompleks. Dalam penelitian perilaku konsumen digital, data mining dapat digunakan untuk mengklasifikasikan minat belanja online berdasarkan atribut-atribut tertentu yang berasal dari konten media sosial[5]. Salah satu algoritma klasifikasi yang banyak digunakan dalam data mining adalah metode Naive Bayes. Metode Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dengan asumsi independensi antar atribut. Meskipun asumsi tersebut bersifat sederhana, metode Naive Bayes terbukti memiliki tingkat akurasi yang baik dan efisiensi komputasi yang tinggi. Oleh karena itu, metode ini sering digunakan dalam berbagai penelitian klasifikasi, termasuk analisis sentimen, prediksi perilaku konsumen, dan pemasaran digital.

Penelitian pertama dilakukan oleh R., Santoso, B., dan Wijaya, A. (2021) yang membahas penerapan algoritma Apriori dalam menganalisis pola pembelian pada data transaksi ritel. Penelitian ini bertujuan untuk menemukan keterkaitan antar produk berdasarkan frekuensi kemunculan item dalam transaksi. Hasil penelitian menunjukkan bahwa algoritma Apriori mampu menghasilkan aturan asosiasi yang bermanfaat dalam pengambilan keputusan bisnis, khususnya dalam pengelolaan stok dan penataan produk. Namun, penelitian ini juga menyimpulkan bahwa proses pembangkitan kandidat itemset menyebabkan waktu komputasi menjadi lebih lama ketika jumlah data transaksi semakin besar [6].

Penelitian kedua dilakukan oleh Jetlie Kang Dkk (2025) yang menerapkan algoritma FP-Growth untuk menganalisis pola penjualan produk elektronik. Penelitian ini berfokus pada efisiensi proses pencarian frequent itemset tanpa melakukan pembangkitan kandidat seperti pada algoritma Apriori. Hasil penelitian menunjukkan bahwa FP-Growth mampu mempercepat proses analisis data serta menghasilkan pola asosiasi yang lebih efektif dalam menentukan kombinasi produk yang sering dibeli secara bersamaan. Penelitian ini membuktikan bahwa FP-Growth lebih sesuai digunakan pada dataset transaksi berukuran besar [7].

Penelitian ketiga dilakukan oleh Gunawan (2022) yang membandingkan kinerja algoritma Apriori dan FP-Growth dalam analisis pola pembelian konsumen pada sistem e-commerce. Penelitian ini menggunakan data transaksi penjualan online sebagai objek penelitian. Hasil pengujian menunjukkan bahwa algoritma FP-Growth memiliki keunggulan dalam hal waktu eksekusi dan efisiensi memori dibandingkan Apriori, terutama pada jumlah data yang besar dan kompleks. Penelitian ini menyimpulkan bahwa FP-Growth lebih optimal digunakan untuk mendukung sistem rekomendasi dan pengelolaan stok produk [8].

Dalam penelitian ini, metode Naive Bayes digunakan untuk mengklasifikasikan minat belanja online pengguna TikTok berdasarkan konten media sosial yang mereka konsumsi. Atribut yang digunakan dalam proses klasifikasi meliputi kualitas konten visual, daya tarik promosi, kredibilitas kreator, serta sentimen komentar. Dengan menghitung probabilitas prior dan probabilitas bersyarat dari setiap atribut, metode Naive Bayes mampu menentukan kelas minat belanja online, yaitu berminat dan tidak berminat. Penelitian ini diharapkan dapat memberikan kontribusi teoretis dalam pengembangan kajian perilaku konsumen digital dan penerapan data mining dalam bidang pemasaran berbasis media sosial[9]. Selain itu, hasil penelitian ini juga diharapkan dapat memberikan kontribusi praktis bagi pelaku usaha dalam merancang strategi konten TikTok yang lebih efektif dan tepat sasaran untuk meningkatkan minat belanja online di era digital.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian ini disusun secara sistematis untuk menentukan minat belanja online menggunakan metode Naive Bayes. Tahapan penelitian meliputi beberapa langkah utama seperti berikut.

a. Identifikasi Masalah

Perkembangan media sosial telah membawa perubahan signifikan dalam pola perilaku konsumen, khususnya dalam aktivitas belanja online [10]. Permasalahan utama yang dihadapi adalah sulitnya mengukur dan mengklasifikasikan minat belanja online pengguna TikTok secara objektif. Minat belanja dipengaruhi oleh berbagai faktor, seperti kualitas visual konten, pesan promosi, kredibilitas kreator, serta tingkat interaksi pengguna yang tercermin dari jumlah like, komentar, dan share.

b. Studi Literatur

Tahap studi literatur dalam penelitian ini berfokus pada tiga konsep utama, yaitu minat belanja online, konten media sosial pada platform TikTok, serta penerapan metode Naive Bayes dalam klasifikasi data. Minat belanja online didefinisikan sebagai kecenderungan atau niat konsumen untuk melakukan pembelian produk atau jasa melalui platform digital. Menurut Kotler dan Keller, minat beli merupakan bagian dari perilaku konsumen yang dipengaruhi oleh persepsi, sikap, dan pengalaman individu terhadap suatu produk [11].

c. Pengumpulan Data

Pengumpulan data merupakan tahap penting dalam penelitian “Klasifikasi Minat Belanja Online pada TikTok Berdasarkan Konten Media Sosial Menggunakan Metode Naive Bayes” karena kualitas data yang diperoleh akan

sangat memengaruhi hasil klasifikasi. Penelitian ini menggunakan dua jenis sumber data, yaitu data primer dan data sekunder, untuk memperoleh informasi yang komprehensif dan relevan dengan tujuan penelitian [12],[13].

d. Penerapan Algoritma (Naïve Bayes)

Penerapan algoritma Naive Bayes dalam penelitian “Klasifikasi Minat Belanja Online pada TikTok Berdasarkan Konten Media Sosial Menggunakan Metode Naive Bayes” bertujuan untuk mengklasifikasikan minat belanja online pengguna ke dalam dua kelas, yaitu berminat dan tidak berminat. Algoritma Naive Bayes dipilih karena kemampuannya dalam melakukan klasifikasi berbasis probabilitas secara sederhana, efisien, dan efektif, khususnya pada dataset dengan jumlah atribut yang cukup banyak [14].

2.1 Minat Belanja Online

Minat belanja online merupakan salah satu konsep penting dalam kajian perilaku konsumen digital. Minat belanja online dapat didefinisikan sebagai kecenderungan psikologis atau niat seseorang untuk melakukan pembelian produk atau jasa melalui platform digital setelah memperoleh informasi dan pengalaman dari media online. Minat ini muncul sebagai hasil dari proses kognitif dan afektif konsumen, yang dipengaruhi oleh persepsi terhadap produk, kepercayaan terhadap penjual, serta pengalaman interaksi dengan media digital. Dalam konteks e-commerce, minat belanja online sering dijadikan indikator awal sebelum terjadinya keputusan pembelian secara aktual. Perkembangan teknologi informasi dan media sosial telah memperluas faktor-faktor yang memengaruhi minat belanja online. Media sosial tidak hanya berfungsi sebagai sarana komunikasi, tetapi juga sebagai media pemasaran yang mampu menyajikan informasi produk secara visual, interaktif, dan persuasif. Pada platform TikTok, konten berbasis video pendek memungkinkan penyampaian pesan promosi yang kreatif dan menarik, sehingga dapat membangun ketertarikan konsumen dalam waktu singkat. Unsur hiburan, keunikan konten, serta keterlibatan emosional pengguna menjadi faktor penting dalam mendorong minat belanja online. Minat belanja online dipengaruhi oleh berbagai faktor, antara lain kualitas informasi produk, daya tarik visual, kredibilitas kreator atau penjual, serta interaksi sosial yang tercermin dari komentar, like, dan share. Interaksi pengguna pada konten TikTok juga dapat membentuk persepsi sosial, di mana respons positif dari pengguna lain dapat meningkatkan kepercayaan dan keyakinan konsumen terhadap produk yang ditawarkan. Selain itu, kemudahan akses dan proses transaksi yang ditawarkan oleh fitur TikTok Shop turut memperkuat minat belanja online pengguna [15],[16].

2.2 Konten Media Sosial

Konten media sosial merupakan strategi komunikasi dan pemasaran digital yang berfungsi untuk menyampaikan pesan, membangun citra, serta memengaruhi perilaku pengguna. Konten media sosial dapat berupa teks, gambar, video, audio, maupun kombinasi dari berbagai format tersebut yang dipublikasikan melalui platform media sosial[17]. Dalam konteks pemasaran digital, konten tidak hanya bertujuan untuk memberikan informasi, tetapi juga untuk menarik perhatian, membangun keterlibatan, dan mendorong respons pengguna terhadap produk atau jasa yang ditawarkan. Pada platform TikTok, konten media sosial didominasi oleh video pendek yang bersifat kreatif, visual, dan interaktif. Karakteristik ini menjadikan TikTok sebagai media yang efektif dalam menyampaikan pesan promosi secara cepat dan menarik. Konten yang disajikan biasanya melibatkan unsur hiburan, storytelling, tren populer, serta penggunaan musik dan efek visual yang khas[18]. Selain itu, peran kreator konten atau influencer juga menjadi faktor penting, karena kredibilitas dan daya tarik kreator dapat meningkatkan kepercayaan pengguna terhadap produk yang dipromosikan. Konten media sosial pada TikTok dapat dianalisis melalui berbagai indikator, seperti kualitas visual video, kejelasan pesan promosi, relevansi konten dengan kebutuhan pengguna, serta intensitas interaksi yang dihasilkan. Interaksi pengguna yang tercermin dari jumlah like, komentar, dan share menunjukkan tingkat keterlibatan audiens terhadap konten yang disajikan. Komentar pengguna juga dapat menggambarkan sentimen positif atau negatif yang berpengaruh terhadap persepsi calon konsumen.

2.3 Algoritma Naïve Bayes

Naive Bayes merupakan salah satu algoritma klasifikasi yang banyak digunakan dalam bidang data mining dan machine learning karena kesederhanaan dan efisiensinya. Metode ini didasarkan pada Teorema Bayes, yang menjelaskan hubungan antara probabilitas suatu kejadian dengan kondisi tertentu[19]. Naive Bayes bekerja dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan nilai atribut yang dimilikinya. Algoritma ini disebut “naive” karena mengasumsikan bahwa setiap atribut bersifat independen satu sama lain, meskipun dalam praktiknya asumsi tersebut sering kali tidak sepenuhnya terpenuhi. Dalam proses klasifikasi, Naive Bayes menghitung probabilitas prior untuk setiap kelas serta probabilitas bersyarat dari setiap atribut terhadap kelas tersebut. Probabilitas prior menggambarkan peluang awal suatu kelas sebelum mempertimbangkan atribut, sedangkan probabilitas bersyarat menunjukkan pengaruh atribut tertentu terhadap kelas yang diamati[20]. Kelas dengan nilai probabilitas posterior tertinggi akan dipilih sebagai hasil klasifikasi. Rumus dasar yang digunakan dalam metode Naive Bayes adalah Teorema Bayes, yaitu perbandingan antara probabilitas bersyarat dan probabilitas total suatu kejadian.

3. HASIL PEMBAHASAN

3.1 Dataset

Fokus perhitungan Naive Bayes dalam penelitian ini menggunakan atribut **Jenis Konten**, karena atribut tersebut paling merepresentasikan karakter konten media sosial dalam memengaruhi minat belanja online.

a. Variabel Kelas (Y)

Minat Belanja:

Tertarik

Tidak Tertarik

b. Variabel Fitur (X)

Data yang tersedia, Perhitungan difokuskan pada :

Jenis Konten

Berikut dapat dilihat pada Tabel 1 merupakan Dataset yg digunakan pada penelitian ini :

Tabel 1. Dataset

No	Caption	Hashtag	Jenis Konten	Minat Belanja
1	Diskon besar sepatu branded hari ini	#diskon #sepatu	Promosi	Tertarik
2	Flash sale tas wanita murah	#flashsale #tas	Promosi	Tertarik
3	Promo beli satu gratis satu kosmetik	#promo #kosmetik	Promosi	Tertarik
4	Live shopping diskon 70 persen	#liveshopping	Promosi	Tertarik
5	Promo skincare remaja	#skincare #promo	Promosi	Tertarik
6	Promo baju muslim terbaru	#bajumuslim #promo	Promosi	Tertarik
7	Diskon besar elektronik	#elektronik #diskon	Promosi	Tertarik
8	Promo sepatu sekolah	#sepatu #promo	Promosi	Tertarik
9	Cerita liburan keluarga	#travel #vlog	Vlog	Tidak Tertarik
10	Cerita pengalaman kerja	#cerita #motivasi	Vlog	Tidak Tertarik
11	Cerita keseharian	#dailyvlog	Vlog	Tidak Tertarik
12	Cerita pengalaman kuliah	#mahasiswa	Vlog	Tidak Tertarik
13	Tips hidup sehat alami	#sehat #motivasi	Edukasi	Tidak Tertarik
14	Edukasi cara menabung	#edukasi #keuangan	Edukasi	Tidak Tertarik
15	Tips diet sehat	#diet #sehat	Edukasi	Tidak Tertarik
16	Edukasi public speaking	#edukasi #tips	Edukasi	Tidak Tertarik
17	Tips mengatur waktu	#produktif	Edukasi	Tidak Tertarik
18	Quotes motivasi harian	#quotes #motivasi	Edukasi	Tidak Tertarik
19	Tips belajar efektif	#belajar #tips	Edukasi	Tidak Tertarik
20	Edukasi menjaga kesehatan	#kesehatan	Edukasi	Tidak Tertarik
21	Edukasi literasi digital	#literasidigital	Edukasi	Tidak Tertarik
22	Tips manajemen stres	#mentalhealth	Edukasi	Tidak Tertarik
23	Tips mengelola keuangan	#keuangan	Edukasi	Tidak Tertarik
24	Video joget tren terbaru	#fyp #dance	Hiburan	Tidak Tertarik
25	Konten lucu kucing	#kucing #lucu	Hiburan	Tidak Tertarik
26	Video prank lucu	#prank #hiburan	Hiburan	Tidak Tertarik
27	Video komedi singkat	#komedi #fyp	Hiburan	Tidak Tertarik
28	Video challenge viral	#challenge #fyp	Hiburan	Tidak Tertarik
29	Video parodi lucu	#parodi #hiburan	Hiburan	Tidak Tertarik
30	Review jujur skincare viral	#skincare #review	Review Produk	Tertarik
31	Unboxing HP terbaru fitur lengkap	#unboxing #gadget	Review Produk	Tertarik
32	Rekomendasi jam tangan murah	#jam #rekomenadasi	Review Produk	Tertarik
33	Review sepatu olahraga original	#sepatu #review	Review Produk	Tertarik
34	Unboxing tas branded	#tas #unboxing	Review Produk	Tertarik
35	Rekomendasi headphone murah	#headphone #gadget	Review Produk	Tertarik
36	Review lipstick tahan lama	#lipstick #review	Review Produk	Tertarik
37	Unboxing laptop murah	#laptop #unboxing	Review Produk	Tertarik
38	Review smartwatch terbaru	#smartwatch #review	Review Produk	Tertarik
39	Review kamera vlog	#kamera #review	Review Produk	Tertarik
40	Unboxing sepatu running	#running #unboxing	Review Produk	Tertarik
41	Review powerbank original	#powerbank #review	Review Produk	Tertarik
42	Unboxing kosmetik lokal	#kosmetik #unboxing	Review Produk	Tertarik
43	Review earphone wireless	#earphone #review	Review Produk	Tertarik
44	Cerita makanan	#makanan	Vlog	Tidak Tertarik
45	Cerita liburan	#liburan	Vlog	Tidak Tertarik
46	Cerita Kegiatan PKL	#mahasiswa	Vlog	Tidak Tertarik
47	Cerita DIY barang aesthetic	#vlog	Vlog	Tidak Tertarik
48	Cerita kejutan	#vlog	Vlog	Tidak Tertarik
49	Cerita makanan unik	#makanan	Vlog	Tidak Tertarik
50	Cerita liburan kampus	#liburan	Vlog	Tidak Tertarik

No	Caption	Hashtag	Jenis Konten	Minat Belanja
51	Cerita memasak	#vlog masak	Vlog	Tidak Tertarik
52	Konten dance santai	#dance #viral	Hiburan	Tidak Tertarik
53	Sketsa komedi ringan	#sketsa #hiburan	Hiburan	Tidak Tertarik
54	Video lucu anak kecil	#lucu #hiburan	Hiburan	Tidak Tertarik
55	Video hiburan keluarga	#keluarga #hiburan	Hiburan	Tidak Tertarik
56	Video lucu hewan	#hewan #lucu	Hiburan	Tidak Tertarik
57	Video hiburan parodi	#parodi	Hiburan	Tidak Tertarik
58	Video lawakan singkat	#lawak	Hiburan	Tidak Tertarik
59	Rekomendasi produk viral TikTok	#produk viral #tiktokshop	Review Produk	Tertarik
60	Review serum wajah	#serum #review	Review Produk	Tertarik
61	Unboxing keyboard mechanical	#keyboard #unboxing	Review Produk	Tertarik
62	Review mouse gaming	#mouse #gaming	Review Produk	Tertarik
63	Unboxing smartwatch anak	#smartwatch #anak	Review Produk	Tertarik
64	Review blender portable	#blender #review	Review Produk	Tertarik
65	Unboxing vacuum cleaner	#vacuum #unboxing	Review Produk	Tertarik
66	Review rice cooker	#ricecooker	Review Produk	Tertarik
67	Unboxing air fryer	#airfryer	Review Produk	Tertarik
68	Review stroller bayi	#stroller	Review Produk	Tertarik
69	Unboxing popok bayi	#popok	Review Produk	Tertarik
70	Edukasi manajemen waktu	#produktifitas	Edukasi	Tidak Tertarik
71	Tips hidup minimalis	#minimalis	Edukasi	Tidak Tertarik
72	Edukasi kebiasaan pagi	#rutinitas	Edukasi	Tidak Tertarik
73	Tips parenting	#parenting	Edukasi	Tidak Tertarik
74	Edukasi perkembangan anak	#tumbuhkembang	Edukasi	Tidak Tertarik
75	Edukasi keselamatan berkendara	#safety	Edukasi	Tidak Tertarik
76	Edukasi perawatan motor	#perawatanmotor	Edukasi	Tidak Tertarik
77	Edukasi tips berkendara	#tipsberkendara	Edukasi	Tidak Tertarik
78	Diskon kosmetik akhir tahun	#kosmetik #diskon	Promosi	Tertarik
79	Promo tas kerja wanita	#tas #promo	Promosi	Tertarik
80	Promo gadget murah	#gadget #promo	Promosi	Tertarik
81	Diskon jam tangan pria	#jam #diskon	Promosi	Tertarik
82	Promo flash sale tengah malam	#flashsale	Promosi	Tertarik
83	Diskon tas sekolah	#tas #diskon	Promosi	Tertarik
84	Promo sepatu lari	#sepatu #promo	Promosi	Tertarik
85	Konten hiburan santai	#hiburan	Hiburan	Tidak Tertarik
86	Video hiburan ringan	#fun	Hiburan	Tidak Tertarik
87	Video lucu keluarga	#family	Hiburan	Tidak Tertarik
88	Video hiburan otomotif	#otomotif	Hiburan	Tidak Tertarik
89	Video hiburan motor	#motor	Hiburan	Tidak Tertarik
90	Video hiburan otomotif lucu	#otomotiflucu	Hiburan	Tidak Tertarik
91	Promo skincare pria	#skincarepria	Promosi	Tertarik
92	Diskon jaket pria	#jaket #diskon	Promosi	Tertarik
93	Promo alat dapur murah	#alatdapur	Promosi	Tertarik
94	Promo peralatan rumah	#perabot	Promosi	Tertarik
95	Promo perlengkapan bayi	#bayi #promo	Promosi	Tertarik
96	Promo mainan anak	#mainananak	Promosi	Tertarik
97	Promo sepeda lipat	#sepeda #promo	Promosi	Tertarik
98	Promo oli kendaraan	#oli	Promosi	Tertarik
99	Promo aksesoris mobil	#aksesoris mobil	Promosi	Tertarik
100	Promo perlengkapan mobil	#mobil #promo	Promosi	Tertarik

3.2 Penerapan Algoritma Naïve Bayes

Adapun Langkah perhitungan dengan menggunakan Naïve Bayes adalah sebagai berikut :

Tahapan Pertama: Melakukan Perhitungan Probabilitas Prior

Probabilitas prior adalah peluang awal suatu kelas tanpa memperhatikan fitur.

a. Prior Kelas Tertarik

$$P(\text{Tertarik}) = \frac{\text{Jumlah Data Tertarik}}{\text{Jumlah Data Tidak Tertarik}}$$

$$P(\text{Tertarik}) = \frac{50}{100} = 0,5$$

b. Prior Kelas Tidak Tertarik

$$P(\text{Tidak Tertarik}) = \frac{50}{100} = 0,5$$

c. Tabulasi Frekuensi

Tabel 2 menyajikan hasil tabulasi frekuensi data yang digunakan dalam penelitian ini. Tabel ini menunjukkan jumlah kemunculan setiap item/variabel dalam keseluruhan data sehingga memberikan gambaran awal mengenai distribusi dan tingkat kemunculan data. Melalui tabulasi frekuensi ini, dapat diketahui item yang paling sering muncul maupun yang memiliki frekuensi rendah, yang selanjutnya menjadi dasar dalam proses analisis data pada tahap berikutnya.

Tabel 2. Tabulasi Frekuensi

No	Jenis Konten	Tertarik	Tidak Tertarik	Total
1	Promosi	25		25
2	Review Produk	25		25
3	Hiburan		19	19
4	Edukasi		19	19
5	Vlog		12	12
	Total	50	50	100

Tahapan kedua : Melakukan Perhitungan Probabilitas Kondisional Untuk Kelas Tertarik

a. Jenis Konten => Promosi

$$P(\text{Promosi} \mid \text{Tertarik}) = \frac{25}{50} = 0,5$$

b. Jenis Konten => Review Produk

$$P(\text{Review Produk} \mid \text{Tertarik}) = \frac{25}{50} = 0,5$$

c. Jenis Konten => Hiburan

$$P(\text{Hiburan} \mid \text{Tertarik}) = \frac{0}{50} = 0$$

d. Jenis Konten => Edukasi

$$P(\text{Edukasi} \mid \text{Tertarik}) = \frac{0}{50} = 0$$

e. Jenis Konten => Vlog

$$P(\text{Vlog} \mid \text{Tertarik}) = \frac{0}{50} = 0$$

Melakukan Perhitungan Probabilitas Kondisional Untuk Kelas Tidak Tertarik

a. Jenis Konten => Promosi

$$P(\text{Promosi} \mid \text{Tertarik}) = \frac{0}{50} = 0$$

b. Jenis Konten => Review Produk

$$P(\text{Review Produk} \mid \text{Tertarik}) = \frac{0}{50} = 0$$

c. Jenis Konten => Hiburan

$$P(\text{Hiburan} \mid \text{Tertarik}) = \frac{19}{50} = 0,38$$

d. Jenis Konten => Edukasi

$$P(\text{Edukasi} \mid \text{Tertarik}) = \frac{19}{50} = 0,38$$

e. Jenis Konten => Vlog

$$P(\text{Vlog} \mid \text{Tertarik}) = \frac{12}{50} = 0,24$$

Melakukan Perhitungan Klasifikasi

a. Kasus 1 Jenis Konten => Promosi

Hitung Posterior Tertarik

$$\begin{aligned}
 &P(\text{Tertarik} \mid \text{Promosi}) \\
 &= P(\text{Promosi} \mid \text{Tertarik}) \times P(\text{Tertarik}) \\
 &= 0,5 \times 0,5 \\
 &= 0,25
 \end{aligned}$$

Hitung Posterior Tidak Tertarik

$$\begin{aligned}
 &P(\text{Tidak Tertarik} \mid \text{Promosi}) \\
 &= P(\text{Promosi} \mid \text{Tidak Tertarik}) \times P(\text{Tidak Tertarik}) \\
 &= 0 \times 0,5 \\
 &= 0 \quad \text{Karena } 0,25 > 0 \text{ maka kelas terpilih Tertarik}
 \end{aligned}$$

- b. Kasus 2 Jenis Konten => Review Produk
Hitungan Posterior Tertarik

$$\begin{aligned}
 &P(\text{Tertarik} \mid \text{Review Produk}) \\
 &= P(\text{Review Produk} \mid \text{Tertarik}) \times P(\text{Tertarik}) \\
 &= 0,5 \times 0,5 \\
 &= 0,25
 \end{aligned}$$

Hitungan Posterior Tidak Tertarik

$$\begin{aligned}
 &P(\text{Tidak Tertarik} \mid \text{Review produk}) \\
 &= P(\text{Review Produk} \mid \text{Tidak Tertarik}) \times P(\text{Tidak Tertarik}) \\
 &= 0 \times 0,5 \\
 &= 0 \quad \text{Karena } 0,25 > 0 \text{ maka kelas terpilih Tertarik}
 \end{aligned}$$

- c. Kasus 3 Jenis Konten => Hiburan
Hitungan Posterior Tertarik

$$\begin{aligned}
 &P(\text{Tertarik} \mid \text{Hiburan}) \\
 &= P(\text{Hiburan} \mid \text{Tertarik}) \times P(\text{Tertarik}) \\
 &= 0 \times 0,5 \\
 &= 0
 \end{aligned}$$

Hitungan Posterior Tidak Tertarik

$$\begin{aligned}
 &P(\text{Tidak Tertarik} \mid \text{Hiburan}) \\
 &= P(\text{Hiburan} \mid \text{Tidak Tertarik}) \times P(\text{Tidak Tertarik}) \\
 &= 0,38 \times 0,5 \\
 &= 0,19 \quad \text{Karena } 0 < 0,19 \text{ maka kelas terpilih Tidak Tertarik}
 \end{aligned}$$

- d. Kasus 4 Jenis Konten => Edukasi
Hitungan Posterior Tertarik

$$\begin{aligned}
 &P(\text{Tertarik} \mid \text{Edukasi}) \\
 &= P(\text{Edukasi} \mid \text{Tertarik}) \times P(\text{Tertarik}) \\
 &= 0 \times 0,5 \\
 &= 0
 \end{aligned}$$

Hitungan Posterior Tidak Tertarik

$$\begin{aligned}
 &P(\text{Tidak Tertarik} \mid \text{Edukasi}) \\
 &= P(\text{Edukasi} \mid \text{Tidak Tertarik}) \times P(\text{Tidak Tertarik}) \\
 &= 0,38 \times 0,5 \\
 &= 0,19 \quad \text{Karena } 0 < 0,19 \text{ maka kelas terpilih Tidak Tertarik}
 \end{aligned}$$

- e. Kasus 5 Jenis Konten => Vlog
Hitungan Posterior Tertarik

$$\begin{aligned}
 &P(\text{Tertarik} \mid \text{Vlog}) \\
 &= P(\text{Vlog} \mid \text{Tertarik}) \times P(\text{Tertarik}) \\
 &= 0 \times 0,5 \\
 &= 0
 \end{aligned}$$

Hitungan Posterior Tidak Tertarik

$$\begin{aligned}
 &P(\text{Tidak Tertarik} \mid \text{Vlog}) \\
 &= P(\text{Vlog} \mid \text{Tidak Tertarik}) \times P(\text{Tidak Tertarik})
 \end{aligned}$$

$$= 0,24 \times 0,5$$

= 0,12 Karena $0 < 0,12$ maka kelas terpilih Tidak Tertarik

Hasil Akhir

Promosi dan Review Produk selalu diklasifikasikan sebagai *Tertarik*

Hiburan, Edukasi, dan Vlog selalu diklasifikasikan sebagai *Tidak Tertarik*

4. KESIMPULAN

Berdasarkan hasil pengolahan dataset sebanyak 100 data, diperoleh distribusi kelas yang seimbang, yaitu 50 data (50%) termasuk dalam kategori Tertarik dan 50 data (50%) termasuk dalam kategori Tidak Tertarik. Nilai probabilitas prior untuk masing-masing kelas adalah $P(\text{Tertarik}) = 0,5$ dan $P(\text{Tidak Tertarik}) = 0,5$, yang menunjukkan bahwa tidak terdapat kecenderungan awal terhadap salah satu kelas. Hasil tabulasi frekuensi menunjukkan bahwa jenis konten Promosi dan Review Produk masing-masing memiliki 25 data (25%) yang seluruhnya diklasifikasikan sebagai Tertarik, sehingga probabilitas kondisional $P(\text{Promosi} | \text{Tertarik}) = 0,5$ dan $P(\text{Review Produk} | \text{Tertarik}) = 0,5$. Sebaliknya, jenis konten Hiburan dan Edukasi masing-masing memiliki 19 data (19%) yang seluruhnya diklasifikasikan sebagai Tidak Tertarik, dengan nilai probabilitas kondisional $P(\text{Hiburan} | \text{Tidak Tertarik}) = 0,38$ dan $P(\text{Edukasi} | \text{Tidak Tertarik}) = 0,38$. Sementara itu, konten Vlog memiliki 12 data (12%) yang seluruhnya termasuk dalam kategori Tidak Tertarik, dengan nilai $P(\text{Vlog} | \text{Tidak Tertarik}) = 0,24$. Hasil perhitungan posterior menunjukkan bahwa untuk konten Promosi dan Review Produk diperoleh nilai $P(\text{Tertarik} | \text{Promosi}) = 0,25$ dan $P(\text{Tertarik} | \text{Review Produk}) = 0,25$, yang lebih besar dibandingkan kelas Tidak Tertarik dengan nilai 0. Sebaliknya, konten Hiburan dan Edukasi menghasilkan nilai $P(\text{Tidak Tertarik} | \text{Hiburan}) = 0,19$ dan $P(\text{Tidak Tertarik} | \text{Edukasi}) = 0,19$, sedangkan konten Vlog menghasilkan nilai $P(\text{Tidak Tertarik} | \text{Vlog}) = 0,12$. Dengan demikian, hasil akhir klasifikasi menunjukkan bahwa 2 dari 5 jenis konten (40%) dikategorikan sebagai Tertarik, sedangkan 3 dari 5 jenis konten (60%) dikategorikan sebagai Tidak Tertarik. Dengan demikian, dapat disimpulkan bahwa klasifikasi minat belanja online berbasis konten TikTok menggunakan metode Naive Bayes mampu memberikan gambaran yang jelas mengenai pengaruh jenis konten terhadap perilaku belanja online, serta memiliki potensi besar untuk dikembangkan lebih lanjut dalam kajian data mining dan pemasaran digital.

REFERENCES

- [1] S. A. S. Mola, Y. C. Luttu, and D. N. Rumlaklak, "Perbandingan Metode Machine Learning dalam Analisis Sentimen Komentar Pengguna Aplikasi InDriver pada Dataset Tidak Seimbang," *J. Sist. Inf. Bisnis*, vol. 14, no. 3, pp. 247–255, 2024, doi: 10.21456/vol14iss3pp247-255.
- [2] F. Saadah, B. U.-S. N. H., "Penerapan Multinomial Naive Bayes untuk Analisis Sentimen Review Skincare Hanasui di TikTok oleh Dokter Detektif," *Prosiding.Uim.Ac.Id*, vol. 11, no. 1, 2025, url: <https://prosiding.uim.ac.id/index.php/sehati/article/view/846>
- [3] M. Andrew, A. Yasin, D. Arman Prasetya, and T. M. Fahrudin, "Analisis Sentimen Tiktok Shop Menggunakan Metode Multinomial Naive Bayes Dan BM25," *J. Ilm. Teknol. Inf. Asia*, vol. 18, no. 02, pp. 24–31, 2024, url: <https://jurnal.asia.ac.id/index.php/jitika/article/view/1004>
- [4] L. Umami, A. Ahmadi, and M. Marhamah, "Pengaruh Iklan Dan Kemudahan Belanja Terhadap Aplikasi Tiktok Terhadap Minat Beli Mahasiswa," *ARMADA J. Penelit. Multidisiplin*, vol. 1, no. 10, pp. 1185–1197, 2023, doi: 10.55681/armada.v1i9.890.
- [5] Magrifatul Zania Maharani, H. Hafiar, and C. Chandratama Priyatna, "Analisis Sentimen Positif Terhadap Brand Kecantikan Avoskin sebagai Eco Friendly Brand di Media Sosial X dan TikTok," *J. Ilm. Mhs.*, vol. 2, no. 2, pp. 25–40, 2024, doi: 10.22373/jim.v2i2.479.
- [6] S. Hermawan and S. Budi, "Analisis dan Prediksi Pertempuran Game Of Thrones Menggunakan Algoritma Random Forest dan Logistic Regression," *J. Strateg. ...*, vol. 3, no. 2, pp. 454–461, 2021, url: <https://mail.strategi.it.maranatha.edu/index.php/strategi/article/view/311>
- [7] P. S. Jetlie Kang, "Analisis Pola Penjualan Produk Elektronik Pada E-Commerce Menggunakan Algoritma FP-Growth," *J. Comasie*, vol. 01, 2025, doi: 10.33884/comasiejournal.v13i1.10267
- [8] R. R. Gunawan bayu atmaja, "Perbandingan Algoritma Apriori dan FP-Growth pada Analisis Perilaku Konsumen Terhadap Pembelian Data Elektronik," *J. Inform. Teknol. dan Sains*, pp. 298–307, 2025, doi: 10.51401/jinteks.v7i1.4850
- [9] B. N. Salsabilah, I. Kadek, and D. Nuryana, "Komparasi Algoritma Naive Bayes dan K Nearest Neighbor dalam Kepuasan Pengguna Fitur Tiktok Shop," *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 4, no. 3, pp. 31–39, 2023., doi: 10.26740/jeisbi.v4i3.54211
- [10] S. C. Nandaresta and C. Warman, "Analisis Sentimen Tanggapan Masyarakat Terhadap Tiktok Shop Dan Shopee Di Twitter Menggunakan Metode Naive Bayes Dan Knn (K- Nearest Neighbor)," *Sismatik*, vol. 12, no. 1, pp. 1–9, 2023, url: <https://sismatik.nusaputra.ac.id/index.php/sismatik/id/article/view/216>
- [11] Meily Anggraini and Feronica Simanjorang, "Efektivitas Online Customer Review Dalam Meningkatkan Minat Beli Konsumen Pada TikTok Shop," *CONTENT J. Commun. Stud.*, vol. 1, no. 02, pp. 10–20, 2023, doi: 10.32734/cjes.v1i02.13227.
- [12] S. Lutfiani, R. Astuti, and Fadhil M Basysyar,M,Kom, "Analisis Sentimen Pengaruh Media Sosial Terhadap Minat Beli Skincare Pada Remaja Di Indonesia Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 2957–2961, 2024, doi: 10.36040/jati.v8i3.9614.
- [13] A. Nurian, M. S. Ma'arif, I. N. Amalia, and C. Rozikin, "Analisis Sentimen Pengguna Aplikasi Shopee Pada Situs Google Play Menggunakan Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, 2024, doi: 10.23960/jitet.v12i1.3631.
- [14] N. Widhiyanta, I. Muhandhis, R. S. Jannah, and L. A. Wulansari, "Skintific Di Tokopedia Menggunakan Support Vector Machine Sentiment Analysis of Skintific Moisturizer Product Reviews," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 18, no. 1, pp. 129–142, 2025, doi: 10.33005/sibc.v18i1.567

- [15] R. Riadi and Mesran, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Analisa Penjualan Parfume," *J. Informatics, Electr. Electron. Eng.*, vol. 2, no. 4, pp. 138–145, 2023, doi: 10.47065/jieec.v2i4.1181.
- [16] T. P. Lestari, "Analisis Text Mining pada Sosial Media Twitter Menggunakan Metode Support Vector Machine (SVM) dan Social Network Analysis (SNA)," *J. Inform. Ekon. Bisnis*, vol. 4, no. 3, pp. 65–71, 2022, doi: 10.37034/infv.v4i3.146.
- [17] H. Hoiriyah, H. Mardiana, M. Walid, and A. K. Darmawan, "Lexicon-Based and Naive Bayes Sentiment Analysis for Recommending the Best Marketplace Selection as a Marketing Strategy for MSMEs," *J. Pilar Nusa Mandiri*, vol. 19, no. 2, pp. 65–76, 2023, doi: 10.33480/pilar.v19i1.4176.
- [18] F. R. Seli and M. A. Ineke Pakereng, "Analisis tingkat kepuasan pengguna Tiktok Shop berdasarkan UI/UX menggunakan metode Naïve Bayes," *IT-Explore J. Penerapan Teknol. Inf. dan Komun.*, vol. 4, no. 2, pp. 211–220, 2025, doi: 10.24246/itexplore.v4i2.2025.pp211-220.
- [19] F. Sulistyo Budi, "Segmentasi Konsumen Tiktok Shop Berdasarkan Perilaku Pembelian Impulsif Menggunakan K-Means Clustering Segmentation of Tiktok Shop Consumers Based on Impulse Buying Behavior Using K-Means Clustering," *Inf. (Jurnal Inform. dan Sist. Informasi)*, vol. 18, no. 2, pp. 1–9, 2025, doi: 10.47701/dutacom.v18i2.5131
- [20] W. R. Wijaya and S. E. Handoyo, "Pengaruh Media Sosial, Kreativitas, Motivasi Terhadap Latar Belakang Pada Era Milenium," *J. Sist. dan Teknol. Inf.*, vol. 05, no. 03, pp. 797–804, 2023, doi: 10.24912/jmk.v5i3.25449