



Prediksi Kepuasan Pelanggan Aplikasi Tiktok dengan Metode Naive Bayes Classifier

Sekar Elok Widhianti ¹, Nurtriana Hidayati ^{*2}

Universitas Semarang

nurtriana.hidayati@gmail.com ^{1,2}

Informasi Artikel

Dikirim :05-01-2025
Direview :10-01-2025
Diterbitkan :30-05-2025

Kata Kunci

Kepuasan Pelanggan,
Naive Bayes, TikTok.

Abstrak

Kegiatan berbelanja secara online telah merajalela dikalangan masyarakat Indonesia. Salah satu platform yang menyediakan layanan belanja online atau e-commerce adalah aplikasi TikTok. TikTok merupakan media sosial untuk menciptakan video dengan tempo pendek yang dilengkapi dengan spesial efek yang mudah digunakan. TikTok menjadi populer dikalangan masyarakat Indonesia setelah terciptanya fitur e-commerce. Namun, kini kompetisi aplikasi penyedia e-commerce makin ketat. Salah satu faktor penting bagi keberlangsungan hidup suatu perusahaan adalah kepuasan pelanggan. Dengan terciptanya kepuasan pelanggan yang baik akan menjadi suatu keuntungan bagi perusahaan tersebut. Maka perlu diadakan penilaian tingkat kepuasan pelanggan pada pengguna fitur e-commerce aplikasi TikTok. Tujuan penelitian ini adalah mengetahui tingkat kepuasan pelanggan dalam berbelanja menggunakan aplikasi TikTok dengan menerapkan data mining pada metode naive bayes. Naive bayes merupakan algoritma data mining yang dapat menghasilkan tingkat akurasi yang baik, Indikator yang digunakan pada penelitian ini adalah biaya, kemudahan akses, kualitas produk dan harga. Data yang digunakan pada penelitian ini berasal dari penyebaran kuesioner terhadap 100 responden yang pernah melakukan transaksi menggunakan aplikasi TikTok. Pengolahan data dilakukan secara urut berdasarkan metode knowledge discovery database dan menggunakan alat bantu RapidMiner. Tingkat akurasi yang dihasilkan dari pengolahan data tersebut adalah 85% dengan perbandingan data training dan data testing sebesar 80:20.

1. PENDAHULUAN

Salah satu aplikasi yang menyediakan layanan e-commerce adalah aplikasi Tiktok. Pada awalnya aplikasi Tiktok merupakan sebuah platform media sosial yang dapat digunakan oleh penggunanya untuk membuat video dengan durasi yang pendek dan dilengkapi oleh spesial efek (Pamungkas & Djulaeka, 2019). Tiktok diluncurkan oleh perusahaan yang bersal dari Tiongkok pada tahun 2016 silam. Kemudian pada Bulan April 2021 secara resmi tiktok meluncurkan fitur terbarunya bernama tiktok shop yang dapat digunakan sebagai media perdagangan secara online. Pada tahun 2022 tercatat bahwa persentase pengguna aplikasi tiktok mengalami peningkatan hingga 10% dari tahun 2021 (Ameliah, Negara et al, 2022).

Platform media sosial tiktok memiliki banyak manfaat bagi para pembisnis online salah satunya guna meningkatkan penjualan produknya. Pembuatan konten video yang menarik akan sangat mempengaruhi tingkat penjualan produk yang sedang dipromosikan khususnya pada media sosial aplikasi TikTok (Supriyanto et al, 2023). Salah satu program aplikasi TikTok yang dapat memberikan keuntungan bagi para pembisnis adalah program pemasaran afiliasi atau yang dikenal TikTok affiliate. Pada program ini para content creator dapat menautkan produk yang sedang dipromosikan dengan video promosi yang telah dibuatnya. Dengan demikian, penyebaran iklan menjadi efisien untuk sampai kepada para calon konsumen.

Selain itu, pemasaran afiliasi juga bersifat menguntungkan bagi para konsumen karena dengan pemasaran ini para konsumen tidak perlu membeli semua varian produk yang sedang dipromosikan. Namun, hal tersebut tidak menjamin rasa puas dari para konsumen. Hal ini dikarenakan banyak faktor yang dapat mempengaruhi kepuasan konsumen. Kepuasan konsumen adalah perasaan yang dirasakan oleh konsumen akibat perbandingan terhadap kinerja suatu produk dengan ekspektasi yang diharapkan (Kotler & Keller, 2016). Konsumen yang merasa puas terhadap produk atau suatu jasa cenderung membeli atau menggunakan kembali jasa tersebut ketika kebutuhan yang sama muncul kembali.

Kepuasan konsumen merupakan faktor penting yang dapat mempengaruhi kelangsungan hidup sebuah perusahaan. Jika suatu perusahaan mampu menciptakan kepuasan konsumen dengan baik, maka hal tersebut dapat menimbulkan keunggulan dalam segi persaingan bisnis bagi perusahaan itu sendiri (Indrasari, 2019). Terciptanya kepuasan konsumen pada aplikasi TikTok akan menimbulkan dampak positif bagi pengguna aplikasi tersebut. Namun, pada saat ini penjual hanya dapat mengetahui kepuasan konsumen berdasarkan komentar dan rating pada produk yang dijualnya saja. Sedangkan, kepuasan konsumen juga dapat dipengaruhi oleh media yang sedang digunakan.

Tujuan pada penelitian ini adalah mengetahui prediksi tingkat kepuasan konsumen dalam berbelanja menggunakan aplikasi TikTok dengan algoritma naïve bayes classifier. Algoritma naïve bayes classifier merupakan algoritma yang dapat memprediksi peluang di masa depan berdasarkan pengalaman pada masa sebelumnya (Jollyta, Ramdhan & Zarlis, 2020). Dengan diketahuinya tingkat kepuasan pelanggan dalam melakukan pembelian menggunakan aplikasi tiktok diharapkan mampu memberikan evaluasi dan pengambilan keputusan terhadap pihak-pihak terkait guna meningkatkan kepuasan pelanggan.

Kerangka Teori

1. Knowledge Discovery Database

KDD atau Knowledge Discovery Database merupakan proses pencarian informasi dan dikembangkan menjadi sebuah pengetahuan yang berguna dan berasal dari himpunan data yang besar (Marisa, Maukar & Akhriza, 2021). KDD sering disamakan dengan data mining. Namun, sebenarnya data mining merupakan salah satu proses yang terdapat di dalam tahap-tahap KDD. Terdapat beberapa proses yang terdapat di dalam KDD diantaranya, yaitu:

- a. *Data Selection*
Kegiatan ini merupakan proses membuang data atau pembersihan data.
- b. *Pre - Processing*
Kegiatan ini merupakan proses membuang data atau pembersihan data.
- c. *Transformasi Data*

Kegiatan ini merupakan proses perubahan data untuk disesuaikan terhadap proses mining.

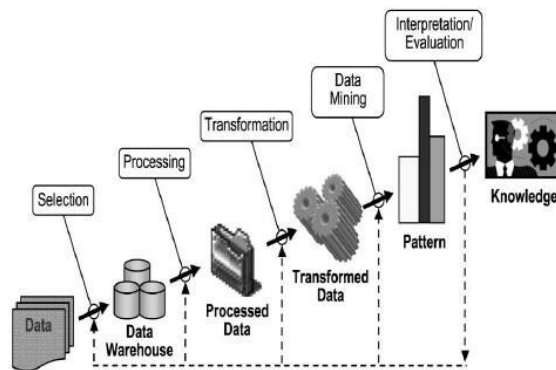
d. Data Mining

Kegiatan ini merupakan mengekstraksi pola yang terdapat di dalam data menggunakan salah satu algoritma yang terdapat pada data mining

e. *Interpretation (Evaluation)*

Kegiatan ini merupakan tahap interpretasi dari sebuah pola menjadi sebuah informasi yang berguna bagi pembaca.

Pada gambar 1 berikut ini merupakan tahap-tahap proses KDD:



Gambar 1. Proses KDD

2. Klasifikasi

Klasifikasi adalah proses pada himpunan data berdasarkan persamaan yang telah ditentukan oleh seorang analis sebelum proses analisis dilaksanakan (Jollyta, Ramdhan & Zarlis, 2020). Tujuan dari proses klasifikasi adalah membentuk sebuah model pembelajaran untuk prediksi di masa mendatang berdasarkan klasifikasi pada data yang dikategorikan menjadi beberapa kelas. Di bawah ini merupakan tahap – tahap dari klasifikasi pada data mining (Jollyta, Ramdhan & Zarlis, 2020):

a. *Learning*

Pada tahap ini dilakukan pencarian pola pada *data training* dan menciptakan deskripsi terhadap model klasifikasi

b. Uji Coba

Model yang telah dibuat selanjutnya diuji pada *data testing* dan data tersebut berbeda dari *data training*

c. *Classification*

Model iterasi yang digunakan untuk membuat klasifikasi pada sisa data dalam *database*.

3. Algoritma Naive Bayes

Naïve bayes merupakan pengklasifikasian dengan teknik peluang dan statistik. Algoritma naïve bayes memiliki asumsi yang kuat akan independensi dari setiap kondisi atau kejadian. Naïve bayes classifier tergolong dalam supervised learning (Amanda, Windarto & Qurniawan). Sehingga pada langkah awal diperlukan data training guna memprediksi suatu kasus. Untuk menggunakan metode naïve bayes diperlukan beberapa peluang sebuah kondisi atau kejadian. Penggunaan algoritma naïve bayes dapat dilihat pada persamaan 1 di bawah ini:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan

$P(H|X)$ = Peluang hipotesis H berdasarkan kondisi X

$P(X|H)$ = Peluang X berdasarkan hipotesis H

$P(H)$ = Peluang hipotesis H

$P(X)$ = Peluang kondisi X

X = Data dengan kelas belum diketahui

H = Hipotesis data dengan suatu kelas spesifik

Dari persamaan (1) di atas dapat disederhanakan menjadi persamaan (2) di bawah ini:

$$P(F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)} \quad (2)$$

Keterangan

C = Mempresentasikan kelas

$F1 \dots Fn$ = Mempresentasikan atribut

Dari persamaan (2) dapat dilihat bahwa semakin banyak atribut yang digunakan, maka akan semakin banyak kompleks pula analisis yang harus dilakukan.

4. Confusion Matrix

Pengukuran performa pada metode naïve bayes classifier dapat dilakukan menggunakan confusion matrix. Confusion matrix dapat diterapkan untuk melakukan pengukuran terhadap hasil klasifikasi yang telah dilakukan dengan sistem terhadap hasil klasifikasi secara aktual. Terdapat 2 jenis klasifikasi, yaitu klasifikasi biner yang hanya dapat melakukan klasifikasi terhadap 2 kelas, dan klasifikasi multiclass yang dapat melakukan klasifikasi terhadap lebih dari 2 kelas (Widodo, 2022). Pada tabel 2.1 berikut ini merupakan representasi dari confusion matrix pada klasifikasi biner:

Tabel 1. Tabel klasifikasi biner

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Hasil representasi dari confusion matrix dapat digunakan sebagai dasar pencarian nilai accuracy, recall dan precision menggunakan persamaan (3), (4) dan (5) berikut ini:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Recall = \frac{TP}{FN+TP} \quad (4)$$

$$Precision = \frac{TP}{FP+TP} \quad (5)$$

Keterangan:

True Positive (TP) : data positif yang diklasifikasikan positif oleh sistem

True Negative (TN) : data negatif yang diklasifikasikan negatif oleh sistem

False Positive (FP) : data negatif yang diklasifikasikan positif oleh sistem

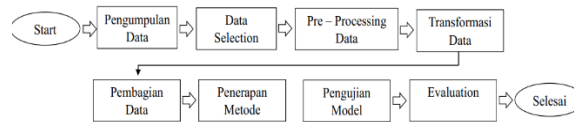
False Negative (FN) : data positif yang diklasifikasikan negatif oleh sistem

5. Split Validation

Split validation adalah proses pembagian data menjadi 2 bagian, yaitu data training dan data testing. Dengan split validation, maka dilakukan eksperimen pada data training berdasarkan rasio pembagian yang telah ditentukan dan sisa dari proses pembagian tersebut dianggap sebagai data testing (Mubarok & Susanti, 2021).

2. METODOLOGI

Pada sebuah penelitian diperlukan tahapan penelitian agar dapat mencapai tujuan dari penelitian tersebut. Gambar 2 berikut ini merupakan tahap-tahap pada penelitian ini:



Gambar 2. Tahap-tahap penelitian

1. Pengumpulan Data

Data diambil dengan penyebaran kuesioner menggunakan alat bantu Google Form kepada pengguna aktif media sosial aplikasi TikTok yang pernah melakukan pembelian barang pada aplikasi tersebut. Total Sampel yang digunakan pada penelitian ini diambil menggunakan rumus lemeshow di bawah ini:

Keterangan:

n = Jumlah sampel

$Z_{1-\alpha/2}^2$ = Tingkat kepercayaan 95%

P = Maksimal estimasi 50%

d^2 = Tingkat kesalahan 10%

$$n = \frac{Z_{1-\alpha/2}^2 P(1-P)}{d^2} \quad (6)$$

$$n = \frac{1,96^2 0.5(1 - 0.5)}{10\%^2}$$

$$n = \frac{3.8416 (0.25)}{0.01}$$

$$n = 96.04 = 100$$

Dari perhitungan yang dilakukan menggunakan persamaan (6) di atas, maka didapatkan hasil sebesar 100 sampel. Sampel tersebut diambil dengan tipe simple random sampling. Simple random sampling merupakan cara pengambilan sampel dari populasi secara acak tanpa memperhatikan strata yang terdapat di dalam populasi tersebut (Rosyidah & Fijra, 2021).

Kuesioner yang digunakan pada penelitian ini berfungsi untuk menilai tingkat kepuasan pelanggan dalam berbelanja menggunakan aplikasi TikTok. Kuesioner tersebut

berisi tentang pertanyaan dan pernyataan mengenai fitur e-commerce berdasarkan variabel biaya, kemudahan akses, kualitas produk dan harga. Tabel 2 berikut ini merupakan definisi operasional pada setiap indikator yang digunakan pada penelitian ini:

Tabel 2. Definisi operasional

Variabel	Definisi Operasional	Indikator
Biaya	Tambahan harga yang ditanggung konsumen dan dirincikan otomatis oleh sistem	1. Biaya Admin 2. Biaya Ongkir
Kemudahan Akses	Kemudahan yang dirasakan konsumen saat menggunakan aplikasi TikTok untuk berbelanja	1. Kejelasan dan kemudahan untuk dipahami 2. Kemudahan untuk digunakan
Kualitas Produk	Kualitas produk yang dirasakan konsumen setelah melakukan pembelian produk pada aplikasi TikTok	1. Daya tahan 2. Kesesuaian 3. <i>Reliabilitas</i> 4. <i>Performance</i>
Harga	Perbandingan harga produk yang dijual pada aplikasi TikTok	1. Keterjangkauan harga 2. Kesesuaian harga dengan kualitas 3. Daya Saing 4. Kesesuaian harga dengan manfaat
Kepuasan Pelanggan	Pengolahan data dengan proses tabulasi	

2. Data Selection

Pada proses data selection pada penelitian ini berfungsi untuk menyeleksi atribut yang akan digunakan pada tahap selanjutnya. Pada tabel 3 di bawah ini merupakan seleksi atribut yang akan dilakukan

Tabel 3. Atribut yang digunakan

Nama Atribut	Keterangan
<i>Timestamp</i>	Tidak digunakan
Nama Responden	Tidak digunakan
Jenis Kelamin	Tidak digunakan
Jawaban indikator 1	Digunakan
Jawaban indikator 2	Digunakan
Jawaban indikator 3	Digunakan
Jawaban indikator 4	Digunakan

3. Pre - Processing

Pada tahap pre – processing ini dilakukan tahap pembuangan instrumen jika terdapat instrumen yang tidak valid dan tidak reliabel. Untuk mengetahui instrumen tersebut valid dan konsisten, maka dilakukan uji validitas dan uji reliabilitas.

a. Uji validitas

Uji validitas merupakan pengukuran dengan tujuan mengetahui kevalidan suatu instrumen. Suatu instrumen dapat dikatakan valid jika instrumen tersebut dapat mengukur yang hendak diukur (Duli, 2019). Teknik yang digunakan pada uji validitas di penelitian ini adalah person product moment. Pada teknik ini instrumen dapat dikatakan valid dan tidak valid jika:

1. r hitung $>$ r tabel (5%), maka valid
2. r hitung $<$ r tabel (5%), maka tidak valid

b. Uji reliabilitas

Uji reliabilitas merupakan alat ukur guna mengetahui kestabilan instrumen apabila dilakukan pengujian berulang kali. Teknik yang digunakan pada uji reliabilitas adalah alpha cronbach. Suatu instrumen dapat dikatakan tidak konsisten jika nilai reliabilitas $<$ 0,6 sedangkan jika nilai reliabilitas $>$ 0,7 dikatakan diterima dan reliabilitas $>$ 0,8 dikatakan baik (Susilowati, 2022).

4. Data Transformasi

Pada tahap ini dilakukan transformasi data yang sesuai untuk dilakukan mining. Pengolahan data pada tahap ini dilakukan dengan menentukan nilai batas interval pada setiap variabel yang digunakan. Pada tabel 4 berikut ini merupakan nilai batas interval yang digunakan:

Tabel 4. Batas nilai interval

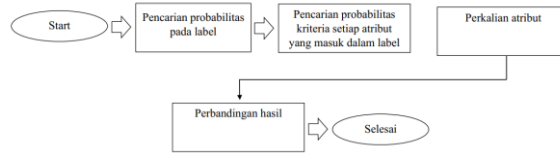
Nama Atribut	Nilai Interval	Keterangan
Biaya	1 - 2	Rendah
	3 - 5	Tinggi
Kemudahan Akses	1 - 2	Rendah
	3 - 5	Tinggi
Kualitas Produk	1 - 2	Rendah
	3 - 5	Tinggi
Harga	1 - 2	Rendah
	3 - 5	Tinggi
Kepuasan Konsumen	16 - 37	Tidak Puas
	38 -59	Cukup Puas
	60 - 80	Sangat Puas

5. Pembagian Data

Data yang telah ditransformasikan dibagi menjadi 2 sebagai data training dan data testing. Rasio perbandingan yang digunakan pada penelitian ini adalah 80:20. Sehingga data ke 1 – 80 merupakan data training, sedangkan data ke 81 – 100 merupakan data testing.

6. Penerapan Metode

Pada tahap ini diterapkan metode naïve bayes classifier. Pada penerapan metode naïve bayes diperlukan beberapa tahap yang harus dilalui. Berikut ini merupakan tahap-tahap metode naïve bayes yang direpresentasikan pada gambar 5 di bawah ini:



Pada tahap ini dilakukan secara urut sesuai dengan tahapan naïve bayes pada gambar 3 di atas.

7. Pengujian Model

Pada tahap ini dilakukan pengujian model menggunakan alat bantu RapidMiner. Pada pengujian model ini pembagian data menggunakan operator split validation dan pembagian secara linear samling. Pembagian data dengan type linear sampling merupakan pembagian yang dilakukan secara urut tanpa mengubah urutan data yang ada (Purnomo et al, 2023).

8. Evaluation

Pada tahap ini menentukan tingkat akurasi dari percobaan yang telah model yang telah dilakukan pada tahap sebelumnya. Penentuan akurasi menggunakan confusion matrix yang dihasilkan secara otomatis pada saat melakukan pengujian model. Dari hasil tersebut ditarik sebuah kesimpulan yang akan menjadi informasi bermanfaat bagi pihak – pihak terkait.

3. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Dari hasil perhitungan menggunakan persamaan (6) dihasilkan total data yang terkumpul sebanyak 100 data. Gambar 6 berikut ini menunjukkan hasil penilaian kepuasan pelanggan terhadap aplikasi TikTok

Tahunnya	Nama Pengguna	Jenis Kelamin	2021.1	2021.2	2021.3	2021.4	2021.1	2021.2	2021.3	2021.4	2021.1	2021.2	2021.3	2021.4	2021.1	2021.2	2021.3	2021.4	
2020-2022 11-44-41	Melissa Dindemariam	Pemrograman	4	4	2	2	4	4	4	2	3	3	3	4	3	4	3	4	3
2020-2022 11-44-51	Yulianawati	Pemrograman	4	4	4	2	3	3	3	3	3	3	4	3	3	4	3	3	3
2020-2022 11-47-51	Agita Andari	Pemrograman	3	3	4	2	3	3	4	2	3	3	3	3	4	3	3	3	3
2020-2022 11-48-01	Harahap Nugroho	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-47-44	Fadhila Egi	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-58-64	Melissa M	Pemrograman	3	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-01-04	Alfetta Anisa Nugroho	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-04-01	Rennan Prita Sembodo	Pemrograman	3	4	3	3	4	4	3	3	3	3	3	4	4	3	4	3	3
2020-2022 11-02-48	Eliana	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-08-09	Agus Triandari	Pemrograman	3	4	3	3	4	4	3	3	3	3	3	4	4	3	4	3	3
2020-2022 11-08-21	Stefaniaty Andarumata	Lain-lain	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-07-01	Yoni	Pemrograman	1	2	2	2	4	4	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-07-01	Siti	Pemrograman	1	2	1	1	4	4	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-08-46	Marissa Rizka Nurfarida Putri	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-08-46	Idara Amalia Lita	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-12-1	Alvin	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-12-1	Alvin	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-12-1	Alvin	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-18-24	Ranita	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-18-24	Ranita	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-20-21	Bayu	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-28-01	Ulu Andri	Lain-lain	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2020-2022 11-28-01	Ulu Andri	Lain-lain	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2020-2022 11-28-01	Ulu Andri	Lain-lain	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2020-2022 11-31-48	Katherine Della	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-04-48	Adina Yulia Satrio	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-58-24	Yuli Nur Nugraha Saputra	Lain-lain	4	4	2	2	4	4	4	4	4	4	4	4	4	4	4	4	4
2020-2022 11-04-01	Rennan Prita Sembodo	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-04-01	Rennan Prita Sembodo	Pemrograman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2020-2022 11-12-21	Ushwanita Cahyaningrum	Pemrograman	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Gambar 6. Hasil penyebaran kuesioner

2. Data Selection

Dilakukan seleksi data yang pada tahap ini, maka dari banyaknya atribut yang terkumpul pada gambar 6 di atas tidak semua atribut digunakan pada tahap selanjutnya. Atribut yang tidak dihilangkan pada tahap ini adalah atribut penilaian kepuasan pelanggan menggunakan aplikasi TikTok.

3. Pre – Processing

Pada tahap ini dilakukan uji instrumen menggunakan alat bantu SPSS versi 26. Uji instrumen dilakukan sebanyak 2 tahap, yaitu uji validitas dan uji reliabilitas. Uji instrumen ini hanya menggunakan 30 data responden pertama. Hasil uji instrumen pada penelitian ini dapat dilihat pada tabel 5 dan tabel 6 di bawah ini:

Tabel 5. Hasil uji validitas

No	Atribut	r hitung	r tabel	Keterangan
1	Biaya	0,864	0,361	Valid
2		0,665	0,361	Valid
3		0,713	0,361	Valid
4		0,865	0,361	Valid
5	Kemudahan	0,904	0,361	Valid
6	Akses	0,612	0,361	Valid
7		0,793	0,361	Valid
8		0,867	0,361	Valid
9	Kualitas	0,946	0,361	Valid
10	Produk	0,879	0,361	Valid
11		0,932	0,361	Valid
12		0,942	0,361	Valid
13		0,765	0,361	Valid
14	Harga	0,904	0,361	Valid
15		0,740	0,361	Valid
16		0,920	0,361	Valid

Tabel 6. Hasil uji reliabilitas

No	Atribut	r hitung	r tabel	Keterangan
1	Biaya	0,928	0,6	Reliabel
2		0,921	0,6	Reliabel
3		0,923	0,6	Reliabel
4		0,925	0,6	Reliabel
5	Kemudahan	0,920	0,6	Reliabel
6	Akses	0,922	0,6	Reliabel
7		0,924	0,6	Reliabel
8		0,916	0,6	Reliabel
9	Kualitas	0,916	0,6	Reliabel
10	Produk	0,920	0,6	Reliabel
11		0,914	0,6	Reliabel
12		0,916	0,6	Reliabel
13		0,919	0,6	Reliabel
14	Harga	0,918	0,6	Reliabel
15		0,924	0,6	Reliabel
16		0,918	0,6	Reliabel

Berdasarkan tabel 5 dan tabel 6 di atas, bahwa semua instrumen yang digunakan pada penelitian ini untuk mengukur kepuasan konsumen dalam menggunakan aplikasi TikTok bersifat valid dan konsisten.

4. Transformasi Data

Pada tahap ini dilakukan transformasi data agar mudah dilakukan proses mining pada data tersebut. Data yang telah terkeumpul di transformasikan berdasarkan batas nilai interval yang telah ditentukan pada tabel 4 di atas. Pada gambar 7 berikut ini merupakan hasil dari transformasi data:

Biaya	Kemudahan Akses	Kualitas Produk	Harga	Kepuasan Konsumen
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Rendah	Tinggi	Tinggi	Tinggi	Cukup Puas
Rendah	Tinggi	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
Rendah	Tinggi	Tinggi	Tinggi	Cukup Puas
Rendah	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Rendah	Tinggi	Tinggi	Cukup Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas

Gambar 7. Hasil transformasi data

5. Pembagian Data

Pada tahap ini merupakan hasil pembagian dataset menjadi data training dan data testing yang dilakukan secara urut. Pada gambar 8 dan gambar 9 berikut ini merupakan hasil pembagian data yang telah dilakukan:

No	Biaya	Kemudahan Akses	Kualitas Produk	Harga	Kepuasan Konsumen
1	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
2	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
3	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
4	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
5	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
6	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
7	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
8	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
9	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
10	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
11	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
12	Rendah	Tinggi	Tinggi	Tinggi	Cukup Puas
13	Rendah	Tinggi	Tinggi	Tinggi	Cukup Puas
14	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
15	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
...
...
77	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
78	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
79	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
80	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas

Gambar 8. Data training

NO	Biaya	Kemudahan Akses	Kualitas Produk	Harga	Kepuasan Konsumen
81	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
82	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
83	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
84	Tinggi	Tinggi	Tinggi	Tinggi	Cukup Puas
85	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
86	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
87	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
88	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
89	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
90	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
91	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
92	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
93	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
94	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
95	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
96	Tinggi	Tinggi	Rendah	Tinggi	Cukup Puas
97	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
98	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas
99	Rendah	Tinggi	Tinggi	Tinggi	Sangat Puas
100	Tinggi	Tinggi	Tinggi	Tinggi	Sangat Puas

Gambar 9. Data testing

6. Penerapan Metode

1. Pencarian probabilitas label

Pada tahap ini dilakukan penerapan metode naïve bayes terhadap data training dan data testing yang telah dibagi pada gambar 8 dan gambar 9. Berikut ini merupakan perhitungan probabilitas pada setiap label berdasarkan pada gambar 8:

a. Sangat Puas

$$\frac{58}{80} = 0,725$$

b. Cukup Puas

$$\frac{20}{80} = 0,25$$

c. Tidak Puas

$$\frac{2}{80} = 0,025$$

2. Pencarian probabilitas kriteria yang masuk dalam label

Pada tahap ini dilakukan pencarian probabilitas kriteria “Tinggi” dan “Rendah” yang masuk ke dalam label “Sangat Puas”, “Cukup Puas” dan “Tidak Puas”. Pada tabel 7, 8 dan 9 berikut ini merupakan perhitungan probabilitas pada setiap kriteria masing-masing indikator yang masuk pada label berdasarkan pada gambar 8:

Tabel 7. Peluang kriteria pada label sangat puas

Atribut	Kriteria	Total	Peluang
Biaya	Tinggi	57	0,9827
	Rendah	1	0,0172
Kemudahan Akses	Tinggi	58	1
	Rendah	0	0
Kualitas Produk	Tinggi	57	0,9827
	Rendah	1	0,0172
Harga	Tinggi	57	0,9827
	Rendah	1	0,0172

Tabel 8. Peluang kriteria pada label cukup puas

Atribut	Kriteria	Total	Peluang
Biaya	Tinggi	16	0,8
	Rendah	4	0,2
Kemudahan Akses	Tinggi	18	0,9
	Rendah	2	0,1
Kualitas Produk	Tinggi	19	0,95
	Rendah	1	0,05
Harga	Tinggi	20	1
	Rendah	0	0

Tabel 9. Peluang kriteria pada label tidak puas

Atribut	Kriteria	Total	Peluang
Biaya	Tinggi	0	0
	Rendah	2	1
Kemudahan Akses	Tinggi	0	0
	Rendah	2	1
Kualitas Produk	Tinggi	0	0
	Rendah	2	1
Harga	Tinggi	0	0
	Rendah	2	1

3. Perkalian pada atribut

Pada tahap ini dilakukan perkalian atribut pada data testing berdasarkan nilai peluang yang telah dihitung pada tabel 7, 8 dan 9. Perkalian atribut ini dilakukan sebanyak 3 kali pada masing-masing label. Berikut ini merupakan perhitungan atribut pada data ke 100 berdasarkan gambar 9:

a. Sangat Puas

$P(\text{Data ke 100}|\text{Sangat Puas})$

$= P(\text{Biaya: Tinggi}|\text{Sangat Puas}) \times P(\text{Kemudahan Akses: Tinggi}|\text{Sangat Puas}) \times P(\text{Kualitas Produk: Tinggi}|\text{Sangat Puas}) \times P(\text{Harga: Tinggi}|\text{Sangat Puas}) \times P(\text{Sangat Puas})$

$= 0,9827 \times 1 \times 0,9827 \times 0,725$

$= 0,6880$

b. Cukup Puas

$P(\text{Data ke 100}|\text{Cukup Puas})$

$= P(\text{Biaya: Tinggi}|\text{Cukup Puas}) \times P(\text{Kemudahan Akses: Tinggi}|\text{Cukup Puas}) \times P(\text{Kualitas Produk: Tinggi}|\text{Cukup Puas}) \times P(\text{Harga: Tinggi}|\text{Cukup Puas}) \times P(\text{Cukup Puas})$

$= 0,8 \times 0,9 \times 0,95 \times 1 \times 0,25$

$= 0,171$

c. Tidak Puas

$$\begin{aligned}
 &= P(\text{Biaya: Tinggi|Tidak Puas}) \times P(\text{Kemudahan Akses: Tinggi|Tidak Puas}) \times P(\text{Kualitas Produk: Tinggi|Tidak Puas}) \times P(\text{Harga: Tinggi|Tidak Puas}) \times P(\text{Tidak Puas}) \\
 &= 0 \times 0 \times 0 \times 0 \times 0,025 \\
 &= 0
 \end{aligned}$$

4. Perbandingan Hasil

Pada tahap ini dilakukan perbandingan hasil perhitungan atribut pada data ke 100 yang sudah dilakukan pada langkah sebelumnya. Untuk mempermudah proses perbandingan hasil maka disajikan dalam tabel 10 berikut ini:

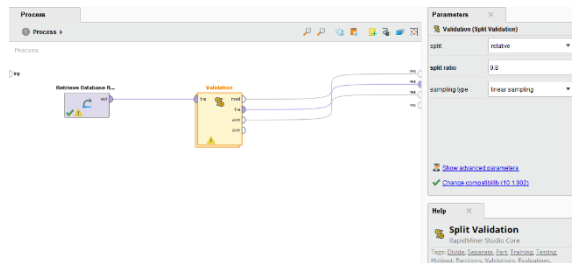
Tabel 10. Peluang kriteria pada label tidak puas

Hasil Label Sangat Puas	Hasil Label Cukup Puas	Hasil Label Tidak Puas
0,6880	0,171	0

Dari perbandingan tabel 10, maka hasil label klasifikasi pada data ke 100 adalah sangat puas.

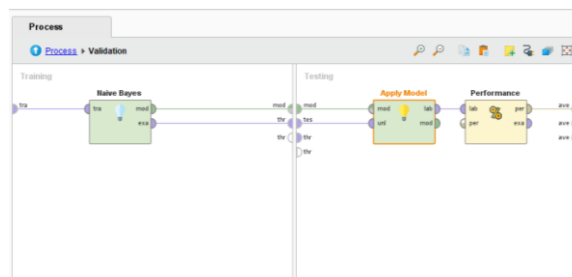
5. Pengujian Model

Pada tahap ini dilakukan pengujian model menggunakan alat bantu RapidMiner. Pada tahap pengujian ini pembagian data menggunakan operator split validation dengan rasio 0,8 untuk data testing dan sisanya digunakan sebagai data training. Tipe sampling yang digunakan pada tahap ini adalah linear sampling. Pada gambar 10 berikut ini merupakan proses pembagian dataset pada RapidMiner:



Gambar 10. Proses pembagian data

Pada tahap selanjutnya dilakukan proses pengujian model dengan operator naïve bayes. Pada gambar 11 berikut ini merupakan proses pengujian model menggunakan operator naïve bayes pada alat bantu RapidMiner:



Gambar 11. Proses pengujian

6. Evaluation

Pada tahap ini dilakukan evaluasi dan penarikan kesimpulan menggunakan confusion matrix yang dihasilkan oleh tahap sebelumnya. Pada gambar 12 berikut ini merupakan hasil akurasi yang dihitung menggunakan confusion matrix yang dilakukan secara otomatis oleh alat bantu RapidMiner:

● Table View ○ Plot View

accuracy: 85.00%

	true Cukup Puas	true Sangat Puas	true Tidak Puas	class precision
pred. Cukup Puas	0	1	0	0.00%
pred. Sangat Puas	2	17	0	89.47%
pred. Tidak Puas	0	0	0	0.00%
class recall	0.00%	94.44%	0.00%	

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TF}{\text{Total seluruh data}} \times 100\% \\
 &= \frac{17 + 0}{20} \times 100\% \\
 &= 85\%
 \end{aligned}$$

4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan pada tahap – tahap sebelumnya didapatkan tingkat akurasi yang dihasilkan dari proses pengujian model adalah 85% dengan prediksi sangat puas yaitu 19 data dan 1 data diprediksi cukup puas. Maka tingkat kepuasan konsumen dalam melakukan pembelian menggunakan aplikasi TikTik mencapai 95% sedangkan 5% lainnya diprediksi cukup puas. Dengan demikian, dapat dikatakan bahwa algoritma naive bayes dapat digunakan untuk memprediksi tingkat kepuasan konsumen dengan cukup baik.

Diharapkan untuk penelitian selanjutnya menggunakan jumlah data yang lebih banyak lagi dan juga menambahkan indikator sebagai alat pengukur kepuasan konsumen. Selain itu, diharapkan juga pada penelitian selanjutnya menggunakan algoritma klasifikasi lainnya dengan tujuan sebagai pembandingan kinerja.

DAFTAR PUSTAKA

- Amanda, A. D., Windarto, A. P., & Qurniawan, H. (2022). Analisis Kepuasan Konsumen Terhadap Pelayanan Store Ms Glow Menggunakan Metode Naive Bayes. *Jurnal Ilmu Komputer*, Vol 1(3): 130-144.
- Ameliah, R., Negara, R. A., Minarto, B., Manurung, T. M., & Akbar, M. (2022). *Status Literasi Digital di Indonesia 2022*. Dipetik Mei 22, 2023, dari [www.kominfo.go.id: https://web.kominfo.go.id/sites/default/files/RepostSurveiStatusLiterasiDigitalIndonesia2022.pdf](https://web.kominfo.go.id/sites/default/files/RepostSurveiStatusLiterasiDigitalIndonesia2022.pdf)
- Duli, N. (2019). *Metodelogi Penelitian Kuantitatif: Beberapa Konsep Dasar untuk Penulisan Skripsi & Analisa Data dengan SPSS*. Yogyakarta: Deepublish.
- Indrasari, M. (2019). *Pemasaran dan Kepuasan Pelanggan*. Surabaya: Unitomo Press.
- Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining dan Penerapan*. Yogyakarta: Deepublish.
- Kotler, P., & Keller, K. L. (2016). *Marketing Management 15 Global Edition*. United States of America: Courier Kendallville.
- Marisa, F., Maukar, A. L., & Akhriza, T. M. (2021). *Data Mining Konsep dan Penereapannya*. Yogyakarta: Deepublish.

- Mubarok, F., & Susanti, S. (2021). Analisa Penjualan Terhadap Caffe Warung Cerbon Menggunakan Klasifikasi Decision Tree. *Jurnal eProsiding Sistem Informasi (POTENSI)*, 2(1), 130-134.
- Pamungkas, R. T., & Djulaeka. (2019). Perlindungan Hukum Pemegang Hak Cipta Atas Lagu Yang Diunggah Pada Aplikasi Tiktok. *Jurnal Simposium Hukum Indonesia*, 1(1), 394-413.
- Purnomo, N. H., Pamungkas, B., & Juliane, C. (2023). Penerapan Algoritma C4.5 Untuk Klasifikasi Tren Pelanggaran Kendaraan Angkutan Barang dengan Metode CRISP-DM. *Jurnal Media Informatika Budidarma*, 7(1), 30-40.
- Rosyidah, M., & Fijra, R. (2022). *Metode Penelitian*. Yogyakarta: Deepublis.
- Widodo, R. B. (2022). *Machine Learning Metode K-Nearest Neighbors Klasifikasi Angka Bahasa Isyarat*. Malang: MNC Publish.
- Supriyanto, A., Chikmah, I. F., Salma, K., & Tamara, A. W. (2023). Penjualan Melalui Tiktok Shop dan Shopee: Menguntungkan yang Mana?. *BUSINESS: Scientific Journal of Business and Entrepreneurship*, 1(1), 1-16.