IDENTIFIKASI KINERJA NAÏVE BAYES UNTUK KLASIFIKASI TINGKAT KEPUASAN PELANGGAN CAFÉ MIE SETANG

Oleh
Ferayanti Boas Gallaran
Universitas Kristen Indonesia Toraja
E-mail: Ferayanti@ukitoraja.ac.id

Article History:
Received: 21-08-2025
Revised: 28-09-2025
Accepted: 24-10-2025

Keywords:
Klasifikasi, Kepuasan pelanggan, Café, Naïve Bayes.

Abstrak: *Dalam era persaingan bisnis yang semakin ketat, industri kuliner dituntut untuk terus berinovasi dan meningkatkan kualitas produk dan layanannya. Salah satu faktor penentu keberhasilan suatu usaha kuliner adalah tingkat kepuasan pelanggan. Penelitian ini bertujuan untuk mengevaluasi kinerja naïve bayes untuk klasifikasi tingkat kepuasan pelanggan Café Mie Setang. Dengan menggunakan data yang diperoleh dari kuesioner yang telah diisi oleh 100 responden. Hasil dari pengujian dataset yang telah dibagi menjadi 70 data latih dan 30 data testing menunjukkan bahwa metode ini mampu memberikan tingkat accuracy 88%, precision 100%, recall 86,95% dan F1-Score 93% dalam mengelompokkan data kepuasan pelanggan. Dari analisis yang dilakukan, diketahui bahwa faktor yang paling berpengaruh terhadap kepuasan pelanggan adalah kualitas pelayanan dan fasilitas. Dengan hasil ini, pihak Cafe dapat menggunakan informasi tersebut untuk memperbaiki dan meningkatkan layanan agar kepuasan pelanggan semakin meningkat. Selain itu, penelitian ini juga membuktikan bahwa metode Naive Bayes cocok digunakan dalam pengolahan data survei pelanggan karena mudah diterapkan dan cukup akurat.*

PENDAHULUAN

Dalam era persaingan bisnis yang semakin ketat, industri kuliner dituntut untuk terus berinovasi dan meningkatkan kualitas produk dan layanannya. Salah satu faktor penentu keberhasilan suatu usaha kuliner adalah tingkat kepuasan pelanggan. Dimasa sekarang dunia usaha mengalami perkembangan yang sangat pesat, sehingga memicu ketatnya persaingan antar sesama pelaku usaha. Kepuasan customer merupakan salah satu faktor terpenting dalam keberlangsungan sebuah perusahaan[1]. Kepuasan pelanggan tidak hanya berdampak pada keberlangsungan bisnis, namun juga dapat menjadi alat pemasaran yang efektif melalui *word of mouth.* Idealnya, sebuah cafe harus memiliki beberapa elemen kunci untuk menciptakan pengalaman yang positif bagi pelanggan. Pertama, kualitas makanan dan minuman harus tinggi, segar, dan lezat, dengan pilihan menu yang variatif untuk menarik lebih banyak pengunjung. Selain itu, pelayanan yang ramah dan efisien sangat penting, staf

yang terlatih dapat memberikan layanan cepat dan sopan. Suasana cafe juga harus nyaman, kebersihan cafe, termasuk area makan dan toilet, harus selalu dijaga untuk menciptakan kesan positif. Lokasi yang strategis dan mudah diakses, serta koneksi internet yang cepat, juga penting untuk kenyamanan pelanggan. Harga yang bersaing, bersama dengan nilai yang diberikan, membuat pelanggan merasa puas. Dengan semua elemen ini, sebuah cafe dapat menciptakan pengalaman yang menyenangkan dan menarik bagi setiap pengunjung.

Café Mie Setang sebagai salah satu industri kuliner, perlu memiliki pemahaman yang mendalam mengenai tingkat kepuasan pelanggannya agar dapat mempertahankan pelanggan yang sudah ada dan menarik pelanggan baru. Cafe Mie Setang sebagai salah satu kuliner yang mengunggulkan berbagai varian Mie dan minuman. Meskipun kafe Mie Setang masih memiliki banyak konsumen setia terutama hari-hari tertentu seperti malam minggu, ada beberapa realita yang terkait dengan klasifikasi kepuasan pelanggan. Pelanggan setia biasanya termasuk dalam kategori puas atau sangat puas karena merasa makanan sesuai dengan selera mereka, harganya terjangkau, atau suasananya nyaman. Namun, ada juga pelanggan netral yang tetap datang hanya karena lokasinya strategis atau tidak ada alternatif lain, meskipun mereka mungkin merasa ada beberapa hal yang bisa diperbaiki. Pelanggan yang tidak puas tetapi tetap datang biasanya didorong oleh kebutuhan tertentu, seperti ketagihan menu khas atau kebiasaan yang sulit ditinggalkan. Dalam hal ini, meskipun Café Mie Setang memiliki pelanggan langganan, tantangan utama adalah memastikan pelanggan puas tetap loyal, meningkatkan pengalaman pelanggan netral, dan mengatasi keluhan pelanggan yang tidak puas agar mereka tetap merasa dihargai. Hal ini penting untuk mempertahankan pelanggan lama sekaligus menarik pelanggan baru. Perlunya mengetahui kepuasan konsumen guna membenahi, mengoreksi baik kesalahan maupun kekurangan yang ada sehingga pelanggan tidak beralih ke tempat lain.

Berdasarkan penelitian yang dilakukan oleh Yudratama dan Epenetus Calvin dengan judul "Klasifikasi Kualitas Kepuasan Pelanggan Pada *Cold N Brew Coffee* Menggunakan Algoritma *Naive Bayes*". Penelitian ini menerapkan metode klasifikasi *Naïve Bayes*, serta menggunakan algoritma *Naïve Bayes.* Berdasarkan hasil pengujian manual yang dilakukan dengan data latih sebanyak 68 data, didapatkan hasil "puas" sebanyak 58 data, dan hasil "tidak puas" sebanyak 10 data. Pengujian data menggunankan metode *Confusion Matrix* pada *rapidMine*r menggunakan algoritma *naïve bayes* menghasilkan tingkat akurasi sebesar 81,25%. Berdasarkan pengolahan data yang sudah dilakukan, maka metode *Naïve Bayes* dapat direkomendasikan untuk memprediksi kualitas kepuasan pelanggan di *Cold N Brew coffee*.[2] Dalam penelitian tersebut ditemukan bahwa metode naïve bayes berhasil digunakan untuk klasifikasi kepuasan pelanggan. Oleh karena itu penulis juga akan mengevaluasi kinerja naïve bayes untuk klasifikasi tingkat menerapkan metode *naïve bayes* untuk mengetahui bagaimana kepuasan pelanggan.

Berdasarkan penelitian tersebut, penulis memiliki gagasan untuk melakukan evaluasi kinerja naïve bayes untuk klasifikasi tingkat kepuasan pelanggan Café Mie Setang. Metode *naïve bayes* akan digunakan untuk mengklasifikasikan berbagai variabel yang dapat mempengaruhi tingkat kepuasan pelanggan cafe Mie Setang seperti, kualitas pelayanan, kualitas produk, harga dan fasilitas. Meski begitu, penggunaan Naive Bayes perlu dievaluasi dari segi akurasi, presisi, recall, dan f1-score untuk mengetahui seberapa baik algoritma ini dalam mengklasifikasikan tingkat kepuasan pelanggan di Café Mie Setang.

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

Berdasarkan hal tersebut, penelitian ini dilakukan untuk mengevaluasi kinerja algoritma Naive Bayes dalam mengklasifikasikan tingkat kepuasan pelanggan Café Mie Setang. Diharapkan hasil dari penelitian ini dapat memberikan rekomendasi yang bermanfaat dalam peningkatan strategi pelayanan dan pengalaman pelanggan. Hasil dari penelitian ini diharapkan dapat memberikan rekomendasi yang lebih spesifik bagi café Mie Setang dalam upaya meningkatkan kualitas produk dan layanannya.

LANDASAN TEORI

Klasifikasi

Sulistyo Basuki mengatakan bahwa klasifikasi berasal dari kata Latin '"*classis*". Klasifikasi adalah proses pengelompokan, artinya mengumpulkan benda/entitas yang sama serta memisahkan benda/entitas yang tidak sama. Secara umum dapat dikatakan bahwa batasan klasifikasi adalah usaha menata alam pengetahuan ke dalam tata urutan sistematis[7]. Towa P. Hmakotrda dan J.N.B. Tairas mengatakan bahwa klasifikasi adalah pengelompokan yang sistematis daripada sejumlah obyek, gagasan, buku atau benda-benda lain ke dalam kelas atau golongan tertentu berdasarkan ciri-ciri yang sama.[8]

Kepuasan Pelanggan

Secara umum, kepuasan pelanggan (*customer satisfaction*) adalah perasaan senang atau kecewa seseorang yang timbul karena membandingkan kinerja yang dipersepsikan produk atau hasil terhadap ekspektasi mereka. Jika kinerja gagal memenuhi ekspektasi, pelanggan akan tidak puas. Jika kinerja sesuai dengan ekspektasi, pelanggan akan puas. Jika kinerja melebihi ekspektasi maka pelanggan akan sangat puas. Kepuasan pelanggan adalah penilaian terhadap produk atau layanan yang memenuhi harapan, yang berkontribusi pada kelangsungan usaha. Manfaatnya termasuk hubungan yang lebih harmonis antara perusahaan dan pelanggan[9]. Beberapa macam faktor-faktor yang dinilai sebagai konsumen untuk tetap bertahan pada salah satu perusahaan, diantaranya:

a. Pelayanan

Pelayanan merupakan pilar utama yang bersifat jasa yang bertujuan memelihara hubungan baik dan meningkatkan hubungan antara produsen dan pelanggan. Ada hal yang tak kalah penting dalam pelayanan yaitu untuk mendengarkan berbagai keluhan atau masukan pelanggan supaya memberi jalan keluar yang tertuju dan produk semakin baik dan diminati para pelanggan.

b. Kualitas Produk

Setelah menampung masukan yang terbaik, kualitas produk sangat penting, karena kualitas produk merupakan representasi (*ekspektasi*) konsumen dengan apa yang produsen buat. Beberapa hal agar kualitas produk dipercaya dengan memperhatikan kesesuain yang dibutuhkan, rentan waktu daya tahan produk yang harus diperhatikan, kelebihan produk yang telah dibuat, dan keandalan suatu produk pembuatan akan mempengaruhi minat konsumen. Dari sinilah kepercayaan (*loyalitas*) datang untuk tetap membeli produk produsen yang dibuat.

c. Harga

Patokan harga menjadi sangat penting di setiap bisnis, karena setiap harga yang ditawarkan perusahaan pada konsumen merupakan tolak ukur produk itu sendiri. Harga akan berpengaruh terhadap permintaan produk pasar. Sebab itu pilihlah harga sebagaimana

tingkat bahan baku, alat yang digunakan dan skill kesulitan menjadikan tolak ukur menjadi nilai harga produk.

d. Promosi dan Strategi Pemasaran yang Tepat

Promosi merupakan bagaimana kita memberi informasi untuk menawarkan, meningkatkan penjualan lebih tinggi dari biasanya dan menciptakan loyalitas suatu produk kepada konsumen. Kegiatan ini berfungsi untuk mempengaruhi konsumen menggunakan jasa atau produk yang akan konsumen merasa puas dengan produk tersebut.

Menurut Tjiptono kepuasan atau ketidakpuasan pelanggan adalah respon pelanggan terhadap evaluasi ketidakpuasan (*disconfirmation*) yang dirasakan antara harapan sebelumnya atau harapan kinerja lainnya dan kinerja aktual produk yang dirasakan setelah memakainya[10].

Cafe

Cafe, yang sering juga disebut *coffeehouse, coffee shop*, atau café, merupakan istilah yang digunakan untuk tempat yang melayani pesanan kopi atau minuman hangat lainnya. Kafe memiliki karakteristik seperti bar atau restoran, tapi berbeda dengan kafetaria. Banyak kafe yang tidak hanya menyediakan kopi, tetapi juga teh bersama dengan makanan ringan, namun ada juga kafe dan resto yang merupakan kafe yang dilengkapi dengan makanan utama (bukan manakanan ringan). Dari sisi budaya, kafe telah tersebar luas untuk menjadi pusat interaksi sosial dimana orang-orang dapat berkumpul, berbicara, menulis, membaca, melawak, atau sekedar mengisi waktu.[3]

Data Mining

Dalam bidang bisnis, teknik data mining digunakan untuk mendukung cakupan yang luas dari aplikasi-aplikasi bisnis *inteligen* seperti customer profiling, targetedmarketing, *workflow management, store layout* dan *fraud detection*. Dalam bidang kedokteran, peneliti dalam bidang biomolekuler dapat menggunakan Teknik data mining untuk menganalisis sejumlah besar data genomic yang sekarang ini telah banyak dikumpulkan untuk menjelaskan struktur dan fungsi gen, memprediksi struktur protein, dan lain-lain.

Naïve Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada *teorema Bayes*. Naïve Bayes merupakan metode klasifikasi probabilitas untuk memprediksi peluang yang akan terjadi berdasarkan pengalaman di masa lalu yang sudah terjadi[4].Ciri utama dari *Naïve Bayes Classifier* ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian.[5]

Naive Bayes

Menurut Olson Delen menjelaskan bahwa *Naïve Bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat kelas keputusan adalah benar, mengingat vektor informasi obyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari "master" tabel keputusan. *Naive Bayes Classifier* bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan oleh Xhemali, Hinde Stone dalam jurnalnya "*Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages*" mengatakan bahwa "*Naïve Bayes Classifier* memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya".

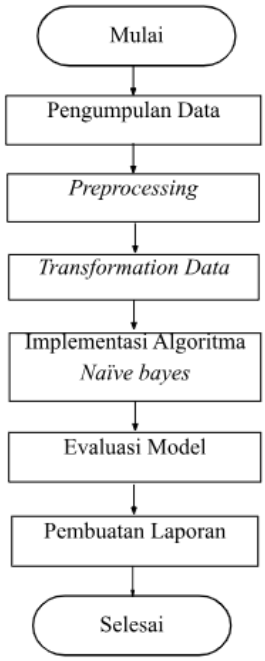
Keuntungan penggunan metode ini, hanya membutuhkan jumlah data pelatihan (training data) yang kecil untuk menentukan estimasi parameter yg diperlukan dalam proses

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

pengklasifikasian, karena yang diasumsikan sebagai variabel independent, maka hanya varian dari suatu variable dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians. penerapan algoritma Naïve Bayes Classifier pada analisis sentiment ulasan pelanggan Mie Gacoan di Twitter pada Oktober 2023. Metode ini dipilih karena kebutuhan data training yang kecil dan sifatnya yang sederhana dan berhasil menghasilkan kinerja yang baik, dengan nilai akurasi sebesar 88,06%, recall sebesar 77,78%, dan presision sebesar 69,23%. Hasil dari penelitian ini dapat memberikan kontribusi positif dalam meningkatkan kualitas produk dan layanan, memperkuat relasi dengan pelanggan, serta mendukung keberlanjutan bisnis di tengah persaingan yang semakin ketat[6].

METODE PENELITIAN

Tahapan penelitian dapat dilihat pada Gambar 3.1 dibawah ini.



Gambar 1

A. Pengumpulan Data

Langkah- langkah pengumpulan data yang akan dilakukan dalam penelitian ini yaitu :

a. Observasi. Pada tahap ini peneliti melakukan observasi langsung pada objek penelitian untuk mendapatkan informasi pendukung terkait penelitian yang dilakukan.

b. Kuisioner. Data yang digunakan dalam penelitian ini adalah data primer. Data ini diperoleh dari hasil kuisioner yang dibagikan kepada pelanggan yang datang ke Café Mie Setang yang berisi pertanyaan mengenai kualitas pelayanan, kualitas produk, harga dan fasilitas. Parameter dari pertanyaan tersebut yaitu sangat puas, puas, cukup puas, tidak puas, dan sangat tidak puas, dengan menggunakan skala likert 1-5 penyebaran kuisioner dilaksanakan pada bulan Maret 2025.

c. Studi Literatur. Pada tahap ini peneliti mengumpulkan informasi dengan membaca referensi berupa jurnal yang berkaitan.

B. Preprocessing

Pada tahapan preprocessing ini peneliti melakukan pengolahan dan penyeleksian data berdasarkan data yang telah dikumpulkan sebelumnya. Hasil dari pengumpulan data mentah hasil kuesioner yang telah di isi oleh pelanggan Café, akan dilakukan pembersihan data. Data akan dibersihkan dari data-data yang tidak diperlukan. Selanjutnya, data hasil kuesioner tersebut disatukan dan dilakukan penginputan kedalam *Microsoft excel,* untuk diberi label untuk masing-masing data sehingga menjadi sebuah dataset.

C. Transformasi Data

Tahap transformasi data dilakukan sesuai dengan jenis data, penulis akan membagi data menjadi data *training* dan data *testing* untuk melatih dan menguji model yang telah dibuat menilai seberapa baik model memprediksi hasil pada data yang belum pernah dilihat, sehingga membantu mengoptimalkan model dan mengidentifikasi kekurangan.

D. Penerapan Metode *Naïve Bayes*

Pada tahapan ini dilakukan klasifikasi menggunakan *software Microsoft excel*. Setelah itu dataset dari *Microsoft excel* di implementasikan dalam Bahasa pemrograman *python* menggunakan *tools Google colab*. Model ini kemudian diuji dengan data pengujian untuk mengevaluasi akurasi dan kemampuan prediksinya. Hasil analisis menunjukkan faktor-faktor yang berpengaruh terhadap kepuasan pelanggan, memberikan wawasan yang dapat digunakan untuk meningkatkan layanan dan produk di cafe.

E. Evaluasi Model

Pada tahap ini, peneliti akan mengukur akurasi model seperti presisi dan recall. Peneliti juga akan menganalisi kesalahan model untuk memahami kekuatan dan kelemahan model. Hasil evaluasi ini digunakan untuk memperbaiki dan mengoptimalkan model sehingga menjadi lebih akurat dan efektif.

F. Pembuatan Laporan

Langkah terakhir yaitu menyimpulkan hasil dari penelitian yang telah dilakukan dan menyusun dalam bentuk laporan.

HASIL DAN PEMBAHASAN

Penelitian ini telah dilakukan dengan tujuan untuk mengevaluasi model naïve bayes dalam mengklasifikasikan tingkat kepuasan pelanggan Café Mie Setang. Hasil penelitian ini akan memberikan gambaran tentang bagaimana pelanggan merasa puas atau tidak puas dengan layanan Café Mie Setang dan bagaiman model *naïve bayes* dapat membantu dalam menganalisis kepuasan pelanggan.

a. Pengumpulan Data

Tahapan awal dalam penelitian ini adalah pengumpulan data. Data yang digunakan dalam penelitian ini merupakan data tingkat kepuasan pelanggan Café Mie Setang. Selanjutnya, peneliti menentukan jumlah responden dengan menggunakan rumus *lemeshow* untuk mengetahui jumlah sampel yang akan digunakan dalam penelitian. Penggunaan rumus *lemeshow* bertujuan untuk membantu melakukan penentuan ukuran sampel secara optimal pada populasi yang tidak diketahui jumlahnya.

Dari perhitungan tersebut diperoleh hasil jumlah sampel minimal yang dibutuhkan dalam penelitian ini adalah 96,04 responden yang kemudian dibulatkan menjadi 100 responden. Dari jumlah responden yang telah didapatkan menggunakan rumus *lemeshow*,

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

selanjutnya jumlah data yang dikumpulkan sebanyak 100 data pelanggan. Data ini mencakup beberapa atribut seperti kualitas pelayanan, kualitas produk, harga dan fasilitas.

b. Uji Validitas dan Reliability

Dalam penelitian ini, uji validitas dan uji reliabilitas digunakan untuk memastikan bahwa instrument pengumpulan data yang digunakan adalah tepat dan konsisten. Uji validitas bertujuan untuk mengevaluasi sejauh mana instrument tersebut mengukur apa yang seharusnya diukur. Sedangkan uji reliabilitas dilakukan untuk menilai konsistensi hasil yang diberikan oleh instrument tersebut dari waktu ke waktu. Adapun tabel distribusi nilai r_{tabel} signifikan 5% dan 1% dapat dilihat pada Tabel 4.2.

Tabel 4. 2 nilai r_{tabel} signifikan 5% dan 1%

N	The Level of Significance		N	The Level of Significance	
	5%	1%		5%	1%
3	0.997	0.999	38	0.320	0.413
4	0.950	0.990	39	0.316	0.408
5	0.878	0.959	40	0.312	0.403
6	0.811	0.917	41	0.308	0.398
7	0.754	0.874	42	0.304	0.393
8	0.707	0.834	43	0.301	0.389
9	0.666	0.798	44	0.297	0.384
10	0.632	0.765	45	0.294	0.380
11	0.602	0.735	46	0.291	0.376
12	0.576	0.708	47	0.288	0.372
13	0.553	0.684	48	0.284	0.368
14	0.532	0.661	49	0.281	0.364
15	0.514	0.641	50	0.279	0.361
16	0.497	0.623	55	0.266	0.345
17	0.482	0.606	60	0.254	0.330
18	0.468	0.590	65	0.244	0.317
19	0.456	0.575	70	0.235	0.306
20	0.444	0.561	75	0.227	0.296
21	0.433	0.549	80	0.220	0.286
22	0.432	0.537	85	0.213	0.278
23	0.413	0.526	90	0.207	0.267
24	0.404	0.515	95	0.202	0.263
25	0.396	0.505	100	0.195	0.256
26	0.388	0.496	125	0.176	0.230
27	0.381	0.487	150	0.159	0.210
28	0.374	0.478	175	0.148	0.194
29	0.367	0.470	200	0.138	0.181
30	0.361	0.463	300	0.113	0.148
31	0.355	0.456	400	0.098	0.128
32	0.349	0.449	500	0.088	0.115

33	0.344	0.442	600	0.080	0.105
34	0.339	0.436	700	0.074	0.097
35	0.334	0.430	800	0.070	0.091
36	0.329	0.424	900	0.065	0.086
37	0.325	0.418	1000	0.062	0.081

a). Uji Validitas

Corelation coefisien yang digunakan untuk uji validitas ini adalah pearson dengan perbandingan nilai rhitung terhadap rtabel . Untuk mendapatkan nilai rtabel dapat dilihat pada tabel distribusi yang terdapat pada tabel 4.2. Hasil nilai rtabel yaitu 0,195 karena jumlah responden yang digunakan adalah 100 orang. Hasil nilai r_{tabel} ini memiliki tingkat kepercayaan 95% dan taraf signifikansi error 5%. Setelah memperoleh nilai rtabel maka langkah berikutnya yaitu membandingkan antara nilai rhitung dengan r_{tabel}. Hasil uji validitas dalam penelitian ini dapat dilihat pada Tabel 4.3 :

Tabel 4. 3 Uji Validitas

Pernyataan	r_{hitung}	Nilai r_{tabel}	Keterangan
Kualitas Pelayanan			
P1	0,830		Valid
P2	0,776		Valid
P3	0,821		Valid
P4	0,638		Valid
P5	0,583	0,195	Valid
Kualitas Produk			
KP1	0,778		Valid
KP2	0,776		Valid
KP3	0,802		Valid
KP4	0,594		Valid
KP5	0,561	0,195	Valid
Harga			
H1	0,730		Valid
H2	0,657		Valid
H3	0,797		Valid
H4	0,664		Valid
H5	0,567	0,195	Valid
Fasilitas			
F1	0,555		Valid
F2	0,638		Valid
F3	0,383		Valid
F4	0,643		Valid
F5	0,706	0,195	Valid

b). Uji Reliabilitas

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

Item pertanyaan dalam variabel dapat dikatakan *reliable* apabila jawaban dari responden adalah konsiten atau stabil dari waktu ke waktu. Item dapat dikatakan *reliable* apabila nilai *cronbach' alpha* ≥ 0,60. Hasil uji reliabilitas dalam penelitian ini dapat dilihat pada Tabel 4.4 :

Tabel 4. 4 Uji Reliabilitas

Pernyataan	Hasil *Crombach's alpha*	Koefisien *Crombach's alpha*	Keterangan
Kualitas Pelayanan	0,785	$\alpha \geq 0{,}60$	Reliabel
Kualitas Produk	0,743	$\alpha \geq 0{,}60$	Reliabel
Harga	0,718	$\alpha \geq 0{,}60$	Reliabel
Fasilitas	0,483	$\alpha \geq 0{,}60$	Reliabel

C. Preprocessing

Pada tahap ini dilakukan proses untuk mengimpor dataset, menghilangkan *noise*, melakukan integrasi data, dan reduksi data sehingga data yang diolah tidak mengalami anomaly yang dapat mempengaruhi hasil analisis klasifikasi kepuasan pelanggan Café Mie Setang. Tujuan dari preprocessing ini adalah untuk memastikan kualitas data yang baik sebelum masuk ke tahap klasifikasi dengan algoritma *Naïve Bayes*. Data yang bersih, bebas dari noise dan *missing values*, serta telah diintegrasikan dengan baik, dapat meningkatkan keakuratan dan interpretabilitas hasil klasifikasi. Tahapan ini merupakan bagian krusial dalam mempersiapkan data sebelum dilakukan klasifikasi lebih lanjut.

Pada proses *preprocessing* dalam penelitian ini, dilakukan beberapa langkah utama:

Import Dataset: Data diimpor dari file Excel yang berisi informasi kepuasan pelanggan Café Mie Setang.

Konversi Data Kategorikal ke Numerik: Beberapa atribut seperti jenis kelamin diubah ke dalam bentuk numerik untuk memudahkan pemrosesan.

Normalisasi Data: Data dinormalisasi untuk memastikan bahwa setiap variabel memiliki bobot yang seimbang dalam model.

Berikut merupakan Langkah-langkah pembuatan :

Import Library yang dibutuhkan

Pada tahap ini, dilakukan pemanggilan berbagai pustaka yang digunakan dalam pengolahan data, pembuatan model, serta evaluasi hasil.

pandas dan numpy: untuk pengolahan data.

sklearn: untuk membagi data, membangun model *Naive Bayes*, dan mengevaluasi performa model.

imblearn: untuk menangani ketidakseimbangan data.

matplotlib dan seaborn: untuk visualisasi data.

google.colab.files: untuk mengunggah dataset di Google Colab.

Tahap Pengunggahan Dataset

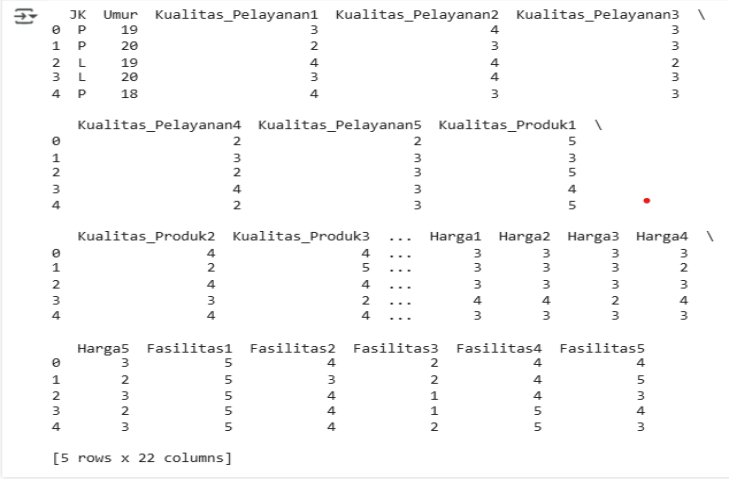
Pada tahap ini, dilakukan pengunggahan file dataset ke dalam lingkungan kerja Google Colab. Pengunggahan ini bertujuan untuk menyediakan data yang akan digunakan dalam proses klasifikasi lebih lanjut, seperti preprocessing, pembuatan model, dan visualisasi data.

Tahap Pengimporan Dataset

Setelah file dataset berhasil diunggah, langkah berikutnya adalah mengimpor dataset ke dalam lingkungan kerja untuk diproses lebih lanjut. Pada tahap ini, digunakan fungsi dari

pustaka pandas untuk membaca file dalam format Excel.

Hasil keluaran dari perintah print(df.head()) dapat dilihat pada Gambar 4.2 berikut:



```
  JK  Umur  Kualitas_Pelayanan1  Kualitas_Pelayanan2  Kualitas_Pelayanan3  \
0  P    19                   3                    4                    3
1  P    20                   2                    3                    3
2  L    19                   4                    4                    2
3  L    20                   3                    4                    3
4  P    18                   4                    3                    3

   Kualitas_Pelayanan4  Kualitas_Pelayanan5  Kualitas_Produk1  \
0                    2                    2                 5
1                    3                    3                 3
2                    2                    3                 5
3                    4                    3                 4
4                    2                    3                 5

   Kualitas_Produk2  Kualitas_Produk3  ...  Harga1  Harga2  Harga3  Harga4  \
0                 4                 4  ...       3       3       3       3
1                 2                 5  ...       3       3       3       2
2                 4                 4  ...       3       3       3       3
3                 3                 2  ...       4       4       2       4
4                 4                 4  ...       3       3       3       3

   Harga5  Fasilitas1  Fasilitas2  Fasilitas3  Fasilitas4  Fasilitas5
0       3           5           4           2           4           4
1       2           5           3           2           4           5
2       3           5           4           1           4           3
3       2           5           4           1           5           4
4       3           5           4           2           5           3

[5 rows x 22 columns]
```

Gambar 4. 2 Tampilan Data saat Berhasil Mengimport Dataset

Tahap perhitungan rata-rata dan kategorisasi kepuasan

Pada tahap ini, dilakukan penghitungan rata-rata skor kepuasan berdasarkan berbagai aspek layanan yang telah dinilai oleh responden. Selanjutnya, hasil rata-rata tersebut digunakan untuk mengkategorikan tingkat kepuasan pelanggan ke dalam dua kategori: "Puas" dan "Tidak Puas". Berikut adalah kode yang digunakan :

```
#Tahap Penghitungan Rata-Rata dan
Kategorisasi Kepuasan

df['Kepuasan_Kategori'] =
df[['Kualitas_Pelayanan1','Kualitas
_Pelayanan2','Kualitas_Pelayanan3',
'Kualitas_Pelayanan4','Kualitas_Pel
```

Penjelasan Proses:

Pada tahap ini, dilakukan dua langkah utama, yaitu menghitung rata-rata skor kepuasan dan mengkategorikan hasilnya. Berikut penjelasan dari masing-masing langkah tersebut:

Menghitung Rata-Rata Skor Kepuasan

Pada langkah pertama, dilakukan penghitungan rata-rata skor kepuasan dari 20 kolom yang mewakili berbagai aspek layanan. Kolom-kolom tersebut mencakup:

Kualitas Pelayanan *(5 atribut)*, Kualitas Produk *(5 atribut)*, Harga (5 atribut), Fasilitas (5 atribut).

Mengkategorikan Kepuasan

Setelah mendapatkan skor rata-rata, langkah berikutnya adalah menentukan kategori kepuasan dengan kriteria berikut:

a. **"Puas"** → jika rata-rata skor lebih dari 3.
b. **"Tidak Puas"** → jika rata-rata skor kurang dari atau sama dengan 3.

Kategori ini dibuat dengan menggunakan fungsi berikut:

```
df['Kepuasan_Kategori'].apply(lambda x: 'Puas' if x > 3 else
```

Fungsi ini akan memeriksa setiap nilai rata-rata di kolom Kepuasan_Kategori dan

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

memberikan label sesuai dengan kriteria yang telah ditentukan.

Memeriksa Hasil Kategorisasi

Untuk memastikan bahwa proses penghitungan rata-rata dan pengelompokan kategori telah berhasil dilakukan. Dataset akan memiliki kolom baru bernama Kepuasan_Kategori yang berisi hasil pengelompokan tingkat kepuasan pelanggan berdasarkan skor rata-rata dari lima dimensi kualitas layanan. Hasil keluaran dataset dapat dilihat dalam Gambar 4.3 berikut :

	Kepuasan_Kategori
0	Puas
1	Puas
2	Puas
3	Puas
4	Puas

Gambar 4. 3 Hasil Perhitungan rata-rata Kategori Kepuasan

Tahap Pemisahan Fitur (X) dan Target (y)

Pada tahap ini, dataset dipisahkan menjadi dua bagian utama: Fitur (X) dan Target (y). Pemisahan ini bertujuan untuk mempersiapkan data sebelum membangun model pembelajaran mesin, di mana:

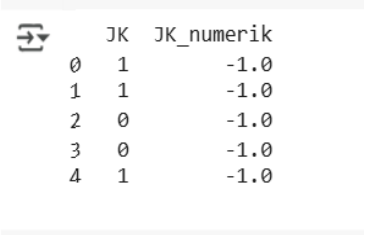
a. Fitur (X) berisi variabel independen yang akan digunakan untuk memprediksi.
b. Target (y) berisi variabel dependen yang ingin diprediksi.

Fitur (X): terdiri dari 20 kolom atribut kepuasan yang akan digunakan sebagai input model.

Target (y): terdiri dari 1 kolom yang berisi kategori kepuasan pelanggan

Tahap Konversi Data Kategorikal menjadi Numerik

Pada tahap ini, dilakukan konversi data kategorikal menjadi numerik. Proses ini diperlukan karena sebagian besar algoritma *machine learning* hanya dapat memproses data dalam bentuk angka (numerik) dan tidak dapat memahami data dalam bentuk kategori atau teks. Hasil keluaran dari proses konversi data kategorikal menjadi numerik dapat dilihat pada Gambar 4.5 berikut :



	JK	JK_numerik
0	1	-1.0
1	1	-1.0
2	0	-1.0
3	0	-1.0
4	1	-1.0

Gambar 4. 5 Hasil Konversi data Kategorikal menjadi Numerik

Tahap Pembagian Data Menjadi Data Latih (*Training*) dan Data Uji (*Testing*)

Pada tahap ini, dataset dibagi menjadi dua bagian, yaitu:

- Data Latih (*Training* Data): Digunakan untuk melatih model agar dapat mengenali pola dari data yang diberikan.
- Data Uji *(Testing* Data): Digunakan untuk menguji kinerja model setelah dilatih, guna mengukur kemampuan model dalam memprediksi data baru yang belum pernah dilihat sebelumnya.

Pembagian data ini dilakukan menggunakan metode train_test_split dari pustaka sklearn.model_selection. Setelah tahap ini selesai, dataset telah terbagi menjadi Data Latih

(*Training* Data): 70% dari dataset, digunakan untuk melatih model, dan Data Uji (*Testing* Data): 30% dari dataset, digunakan untuk menguji performa model. Hasil keluaran dari proses pembagian data menjadi data latih dan data uji dapat dilihat pada Gambar 4.6 berikut :

```
# Bagi Data menjadi Training dan Testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
print("Bentuk data latih dan uji:")
print(X_train.shape, X_test.shape, y_train.shape, y_test.shape)
```

```
Bentuk data latih dan uji:
(70, 22) (30, 22) (70,) (30,)
```

Gambar 4. 6 Hasil Pembagian Data Menjadi Data Latih *(Training)* dan Data Uji *(Testing)*

Tahap *Oversampling* SMOTE untuk Menyeimbangkan Data

Pada tahap ini, dilakukan proses *undersampling* untuk menyeimbangkan jumlah data antar kelas dalam dataset. *Undersampling* adalah teknik dalam *machine learning* yang bertujuan untuk mengurangi jumlah sampel dari kelas mayoritas agar seimbang dengan kelas minoritas.

Tahap Pembuatan dan Pelatihan Model *Naive Bayes*

Pada tahap ini, dilakukan proses pembuatan dan pelatihan model *Naive Bayes* menggunakan algoritma *Gaussian Naive Bayes. Naive Bayes* adalah algoritma machine learning berbasis probabilitas yang didasarkan pada *Teorema Bayes* dengan asumsi bahwa setiap fitur bersifat independen *(naive*). Dalam penelitian ini, digunakan algoritma *Gaussian Naive Bayes* yang mengasumsikan bahwa distribusi data mengikuti distribusi normal (*Gaussian*). Proses pembuatan dan pelatihan model dilakukan melalui beberapa langkah berikut:

[1] Inisialisasi Model Gaussian Naive Bayes

[2] Melatih Model Menggunakan Data Latih

Teorema Bayes dengan rumus:

$$P(X) = \frac{P(X)|C\ x\ P(C)}{P(X)}$$

[3] Melakukan Prediksi pada Data Uji

Setelah tahap ini selesai, model *Naive Bayes* telah berhasil dibangun dan dilatih menggunakan data latih hasil *undersampling*. Proses pembuatan dan pelatihan model serta hasil prediksi pada data uji dapat dilihat pada Gambar 4.8 berikut ;

```
#  Buat dan Latih Model Naive Bayes
model = GaussianNB()
model.fit(X_train_resampled, y_train_resampled)
y_pred = model.predict(X_test)
```

Gambar 4. 8 Hasil Pembuatan dan Pelatihan Model *Naive Bayes*

Tahap Evaluasi Model *Naive Bayes*

Setelah model *Naive Bayes* selesai dilatih dan menghasilkan prediksi, tahap selanjutnya adalah mengevaluasi kinerja model. Evaluasi model bertujuan untuk mengukur seberapa baik model dalam memprediksi label target pada data uji yang belum pernah dilihat

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)

sebelumnya. Dalam penelitian ini, digunakan metrik akurasi dan laporan klasifikasi untuk menilai performa model.

Menghitung Akurasi Model

accuracy_score adalah fungsi dari pustaka *sklearn.metrics* yang digunakan untuk menghitung akurasi model.

Akurasi menunjukkan persentase prediksi yang benar terhadap seluruh data uji.

Menampilkan Hasil Akurasi Model

Menampilkan Laporan Klasifikasi (*Classification Report).*

1. *Precision* (Presisi): Proporsi prediksi benar pada setiap kelas.
2. *Recall* (*Recall*): Proporsi data positif yang berhasil diklasifikasikan dengan benar.
3. *F1-Score*: Rata-rata harmonik dari *precision* dan *recall.*
4. *Support*: Jumlah sampel pada setiap kelas.

Interpretasi Hasil Evaluasi

1) Akurasi: Menunjukkan persentase prediksi benar. Misalnya, 86% berarti dari 200 data uji, 172 di antaranya diprediksi dengan benar.

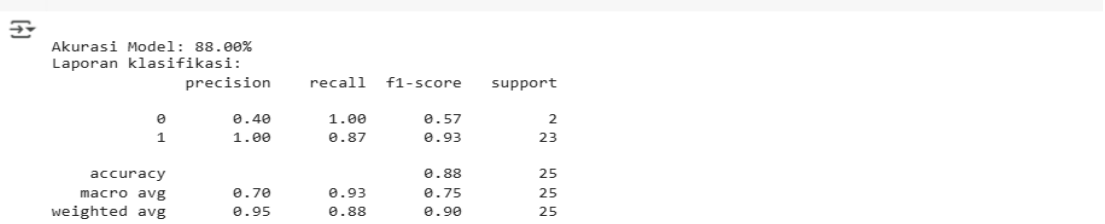
2) *Precision*: Dari seluruh prediksi positif, berapa yang benar? Semakin tinggi nilainya, semakin baik model dalam meminimalkan false positive.

3) *Recall*: Dari seluruh data positif, berapa yang terprediksi dengan benar? Nilai tinggi menunjukkan model mampu mendeteksi kelas positif dengan baik.

4) *F1-Score*: Metrik gabungan yang menunjukkan keseimbangan antara precision dan recall.

5) *Support:* Jumlah sampel pada masing-masing kelas dalam data uji.

Setelah tahap ini selesai, model telah berhasil dievaluasi menggunakan metrik akurasi dan laporan klasifikasi. Hasil akurasi dan performa model berdasarkan hasil evaluasi yang telah dilakukan dapat dilihat pada Gambar 4.9 berikut :

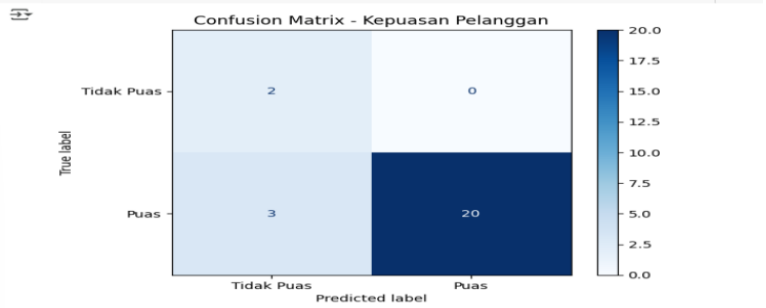


```
Akurasi Model: 88.00%
Laporan klasifikasi:
              precision    recall  f1-score   support

           0       0.40      1.00      0.57         2
           1       1.00      0.87      0.93        23

    accuracy                           0.88        25
   macro avg       0.70      0.93      0.75        25
weighted avg       0.95      0.88      0.90        25
```

Gambar 4. 9 Hasil Evaluasi Model *Naive Bayes*

Tahap Visualisasi *Confusion Matrix*

Pada tahap ini, dilakukan visualisasi hasil prediksi model dengan menggunakan *Confusion Matrix*. *Confusion Matrix* merupakan alat evaluasi yang memvisualisasikan performa model klasifikasi dengan membandingkan hasil prediksi model dan label sebenarnya. Matriks ini membantu dalam memahami kesalahan prediksi model secara lebih rinci dengan menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas.



Gambar 4. 10 Visualisasi *Confusion Matrix*

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, klasifikasi tingkat kepuasan pelanggan di Café Mie Setang berhasil dilaksanakan dengan baik dan metode *Naive Bayes* berhasil digunakan untuk mengklasifikasikan tingkat kepuasan pelanggan di Cafe Mie Setang. Dengan evaluasi metrik yang diperoleh, model ini menunjukkan performa yang baik dalam mengklasifikasikan kepuasan pelanggan dengan tingkat akurasi 88.00%, presisi 100% , *recall* 86,95% dan *F1-Score* 93%. Hasil klasifikasi menunjukkan bahwa lebih banyak pelanggan yang merasa puas dibandingkan dengan yang merasa tidak puas, dengan distribusi prediksi 86 pelanggan puas dan 14 pelanggan tidak puas. Klasifikasi jumlah pelanggan puas dan tidak puas berdasarkan jenis kelamin, jumlah laki-laki yang puas 38 pelanggan dan tidak puas 5 pelanggan, jumlah perempuan yang puas 48 pelanggan dan tidak puas 9 pelanggan.

Berdasarkan hasil yang diperoleh, peneliti merekomendasikan Café Mie Setang untuk meningkatkan kualitas pelayanan dan fasilitas agar pelanggan merasa lebih nyaman dan puas. Beberapa pelanggan merasa bahwa pelayan belum ramah, sopan, serta belum menunjukkan sikap profesional. Selain itu, penampilan pelayan dinilai kurang rapi. Oleh karena itu, disarankan agar pihak café memberikan pelatihan kepada pelayan tentang cara melayani pelanggan dengan baik, serta menetapkan aturan mengenai kerapian dan kebersihan penampilan. Dari sisi fasilitas, toilet dinilai bersih namun terlalu sempit, sehingga membuat pelanggan merasa tidak nyaman. Tempat parkir juga dianggap sempit dan kurang aman. Untuk mengatasi hal ini, café sebaiknya meningkatkan kebersihan toilet secara rutin dan, jika memungkinkan, melakukan sedikit renovasi agar lebih nyaman. Selain itu, pengaturan parkir perlu diperbaiki atau bisa juga bekerja sama dengan pemilik lahan di sekitar café agar parkir lebih luas dan tertata. Dengan perbaikan pada aspek pelayanan dan fasilitas ini, diharapkan kepuasan pelanggan dapat meningkat.

Berdasarkan evaluasi metrik, model ini dapat digunakan sebagai alat analisis untuk membantu peningkatan kualitas pelanggan Café Mie Setang.

DAFTAR PUSTAKA

[1] R. Alfiani and Y. Umaidah, "Prediksi Kepuasan Customer Terhadap Performance Terapis Baby Massage Menggunakan Algoritma Naive Bayes," *BINA INSANI ICT J.*, vol. 9, no. 1, p. 83, Jul. 2022, doi: 10.51211/biict.v9i1.1794.

[2] Calvin and A. Nugroho, "Klasifikasi Kualitas Kepuasan Pelanggan Pada Cold N Brew Coffee Menggunakan Algoritma Naive Bayes," *J. Pendidik. Teknol. Inf. JUKANTI*, vol. 7, no.

2, pp. 189–199, Nov. 2024, doi: 10.37792/jukanti.v7i2.1293.
[3] G. Sukarno and L. Nirawati, "Kontribusi Human Capital Dan Customer Capital Dalam Menggapai Kinerja Café Dan Resto Di Surabaya," *J. IlmU Sos.*, vol. 15, no. 2, p. 137, Nov. 2016, doi: 10.14710/jis.15.2.2016.137-149.
[4] Institut Teknologi Telkom Purwokerto, R. D. Prasetya, R. I. Pasya, and R. A. Bagaskoro, "Memprediksi Tingkat Kepuasan Pelanggan Terhadap Kualitas Suatu Toko Menggunakan Metode Naïve Bayes," *Method. J. Tek. Inform. Dan Sist. Inf.*, vol. 9, no. 2, pp. 18–22, Sep. 2023, doi: 10.46880/mtk.v9i2.1943.
[5] Syarifah, "Pemanfaatan Naïve Bayes Untuk Merespon Emosi Dari Kalimat Berbahasa Indonesia," 2015.
[6] G. Shalihah, R. Kurniawan, and T. Suprapti, "Analisis Sentimen Ulasan Pelanggan Mie Gacoan Menggunakan Metode Naïve Bayes Classifier," *JATI J. Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 593–601, Feb. 2024, doi: 10.36040/jati.v8i1.8302.
[7] O. Widodo, "Mengklasifikasi Dan Menentukan Tajuk Subjek Bahan Perpustakaan".
[8] A. R. Shalaby, M. Hum, and A. Wahidi, "Pengklasifikasian Dan Penataan Ulang Buku Di Perpustakaan Sdn 15 Gelumbang," vol. 1, 2019.
[9] T. W. Widyaningsih, "Implementasi Algoritma Naive Bayes Untuk Analisis Kepuasan Pelanggan Pada Sistem Layanan Jasa Internet (Studi Kasus PT. Kawan Lama Kreatif)," vol. 12, no. 2, 2024.
[10] Gofur, "Pengaruh Kualitas Pelayanan Dan Harga Terhadap Kepuasan Pelanggan," vol. 4, no. 1.

HALAMAN INI SENGAJA DIKOSONGKAN

ISSN 2798-3471 (Cetak)
ISSN 2798-3641 (Online)