J-CEKI: Jurnal Cendekia Ilmiah
Vol. 4, No. 6, Oktober 2025

Penerapan Data Mining untuk Prediksi Kepuasan Pelanggan UKM Warung Kopi Menggunakan Metode Naive Bayes

Cici Yolanda Tanjung[1], Rasiban[2]
[1,2]Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika
E-mail: tanjungciciyolanda@gmail.com[1], rasiban.stikom@yahoo.com[2]

Article History:
Received: 05 Oktober 2025
Revised: 22 Oktober 2025
Accepted: 25 Oktober 2025

Keywords: *kepuasan pelanggan, warung kopi, naive bayes, data mining*

Abstract: *Kepuasan pelanggan merupakan faktor krusial dalam menjaga keberlanjutan bisnis, terutama bagi Usaha Kecil dan Menengah (UKM) seperti warung kopi yang bergantung pada loyalitas pelanggan. Namun, banyak UKM masih menggunakan pendekatan manual dalam menganalisis survei, yang kurang efektif dalam menghasilkan informasi prediktif. Penelitian ini bertujuan untuk mengembangkan model prediksi kepuasan pelanggan menggunakan algoritma Naive Bayes, dengan studi kasus pada sebuah warung kopi di Taman Burung Waduk Pluit, Jakarta Utara. Data dikumpulkan melalui kuesioner yang mencakup lima aspek: kualitas produk, layanan, suasana, harga, dan kebersihan. Algoritma Naive Bayes dipilih karena efisiensinya dalam pemrosesan data dan kemampuannya menghasilkan klasifikasi yang akurat. Kinerja model dibandingkan dengan algoritma Decision Tree dan Regresi Logistik. Hasil menunjukkan bahwa model Naive Bayes mencapai akurasi sebesar 96,05%, serta berhasil mengidentifikasi variabel yang paling berpengaruh terhadap kepuasan pelanggan. Temuan ini memberikan kontribusi penting dalam penerapan pengambilan keputusan berbasis data di sektor UKM, khususnya dalam meningkatkan kualitas layanan dan loyalitas pelanggan.*

PENDAHULUAN

Kepuasan pelanggan merupakan faktor kunci dalam kesuksesan warung kopi skala kecil dan menengah. Hal ini berdampak langsung terhadap loyalitas pelanggan, citra usaha, serta keberlanjutan operasional. Namun, kualitas layanan yang esensial ini sering kali diabaikan oleh pelaku usaha, meskipun terbukti memiliki dampak negatif apabila tidak dikelola secara baik (Permatasari & Susanto, 2022; Adiwijaya & Sari, 2023).

Di Indonesia, kedai kopi skala kecil dan menengah menunjukkan pertumbuhan pesat. Ribuan unit usaha berkontribusi signifikan terhadap perekonomian nasional serta penciptaan lapangan kerja (BPS, 2024; Kementerian Koperasi dan UKM, 2024). Di Kecamatan Penjaringan, Jakarta Utara, fenomena budaya *ngopi* di kalangan Generasi Z telah menciptakan persaingan yang

sangat ketat. Oleh karena itu, peningkatan kualitas layanan menjadi keharusan agar pelaku usaha tetap bertahan.

Sayangnya, sekitar 85% pemilik warung kopi di Penjaringan masih mengalami kesulitan dalam menganalisis hasil survei kepuasan pelanggan. Proses pengolahan data kuesioner masih dilakukan secara manual, tanpa pendekatan analitik yang dapat mengidentifikasi tren atau faktor dominan yang memengaruhi kepuasan pelanggan (Nugroho & Wulandari, 2023).

Penelitian sebelumnya telah membuktikan bahwa algoritma Naive Bayes cukup efektif dalam memprediksi kepuasan pelanggan di berbagai sektor. Pratiwi dan Hendry (2022), misalnya, berhasil menerapkan metode Naive Bayes untuk memprediksi kepuasan pelanggan Shopee dan memperoleh hasil klasifikasi yang akurat. Penelitian lain oleh Amanda et al. (2022) juga menunjukkan akurasi tinggi sebesar 80% dalam menganalisis kepuasan pelanggan pada Store MS Glow. Proses pemodelan pada penelitian-penelitian tersebut dilakukan menggunakan perangkat lunak RapidMiner versi 10.3, yang terbukti memudahkan visualisasi hasil klasifikasi serta mempercepat pemrosesan data. Selain itu, Pradita dan Rasiban menggunakan pendekatan regresi linier dalam prediksi penjualan, yang mengindikasikan potensi analisis data dalam mendukung keputusan strategis.

Dari berbagai studi tersebut dapat disimpulkan bahwa algoritma Naive Bayes memiliki performa klasifikasi yang cukup tinggi. Keunggulan utamanya terletak pada kesederhanaan proses perhitungan, kecepatan dalam membangun model, serta kemampuannya menghasilkan klasifikasi yang akurat, bahkan dalam skala data yang besar. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melanjutkan dan mengembangkan penerapan algoritma Naive Bayes dalam konteks UKM warung kopi, guna menghasilkan model prediksi kepuasan pelanggan yang efisien dan mendukung pengambilan keputusan berbasis data.

LANDASAN TEORI

Data Mining

Menurut Han dan Kamber (2012), *data mining* adalah proses menemukan pola atau pengetahuan dari kumpulan data yang besar dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. Turban et al. (2011) menambahkan bahwa *data mining* memungkinkan peneliti mengidentifikasi hubungan yang tersembunyi dan tren yang signifikan dari data yang ada. Dalam konteks UKM, *data mining* dapat membantu pemilik usaha memahami perilaku pelanggan dan mendukung pengambilan keputusan strategis (Larose, 2014).

Algoritma Naive Bayes

Naive Bayes adalah algoritma klasifikasi berbasis probabilitas yang didasarkan pada Teorema Bayes dengan asumsi independensi antar atribut (Manning et al., 2008). Menurut Zhang (2004), metode ini sederhana namun efektif untuk mengolah data berukuran besar maupun kecil, dengan kecepatan komputasi yang tinggi. Keunggulan Naive Bayes terletak pada kemampuannya memberikan hasil klasifikasi yang baik meskipun data yang digunakan memiliki jumlah sampel terbatas (Sutanto, 2022).

Kepuasan Pelanggan

Kotler dan Keller (2016) mendefinisikan kepuasan pelanggan sebagai tingkat perasaan seseorang setelah membandingkan kinerja atau hasil yang dirasakan dengan harapannya. Menurut Oliver (2010), kepuasan pelanggan berperan penting dalam membentuk loyalitas dan mempertahankan hubungan jangka panjang. Dalam sektor UKM seperti warung kopi, kepuasan

pelanggan mencakup aspek kualitas produk, layanan, suasana, harga, dan kebersihan (Tjiptono, 2014).

Penerapan Data Mining dan Naive Bayes dalam Prediksi Kepuasan Pelanggan

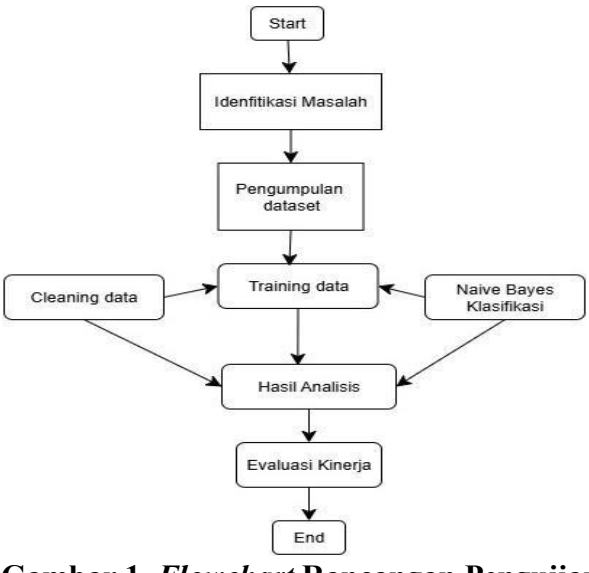
Beberapa penelitian sebelumnya menunjukkan keberhasilan Naive Bayes dalam memprediksi kepuasan pelanggan. Pratiwi dan Hendry (2022) memperoleh akurasi sebesar 88% dalam memprediksi kepuasan pelanggan Shopee, sementara Amanda et al. (2022) mendapatkan akurasi 80% pada analisis kepuasan pelanggan Store MS Glow. Penelitian ini mengadopsi pendekatan serupa dengan fokus pada UKM warung kopi, di mana data primer dikumpulkan melalui kuesioner dan dianalisis menggunakan perangkat lunak RapidMiner versi 10.3.

METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen untuk membangun model prediksi kepuasan pelanggan pada UKM warung kopi menggunakan algoritma Naive Bayes. Data penelitian diperoleh dari kuesioner yang disebarkan kepada pelanggan warung kopi di Taman Burung Waduk Pluit, Jakarta Utara. Variabel yang diukur meliputi kualitas produk, pelayanan, suasana, harga, dan kebersihan. Tahapan penelitian mengikuti kerangka CRISP-DM (*Cross Industry Standard Process for Data Mining*), yang terdiri dari:

1. *Business Understanding* – mendefinisikan tujuan dan permasalahan penelitian.
2. *Data Understanding* – mengumpulkan dan memeriksa data responden.
3. *Data Preparation* – membersihkan data, mengatasi data kosong, dan melakukan transformasi variabel.
4. *Modeling* – menerapkan algoritma Naive Bayes untuk klasifikasi kepuasan pelanggan.
5. *Evaluation* – mengukur akurasi model dan membandingkannya dengan metode *Decision Tree* dan *Logistic Regression.*

Teknik analisis data yang digunakan meliputi analisis deskriptif untuk profil responden, *preprocessing* data (data cleaning dan transformation), penerapan algoritma Naive Bayes, serta evaluasi model menggunakan *confusion matrix* dengan metrik akurasi, presisi, recall, dan F1-score. Alur proses penelitian ditunjukkan pada Gambar 1.



Gambar 1. *Flowchart* Rancangan Pengujian

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan memprediksi kepuasan pelanggan UKM warung kopi di area Taman Burung Waduk Pluit menggunakan algoritma Naive Bayes. Data dikumpulkan melalui kuesioner yang mencakup variabel kualitas produk, pelayanan, suasana, harga, dan kebersihan. Setelah proses pengolahan data, jumlah data valid yang digunakan untuk analisis adalah 76 responden. Proses analisis dilakukan dengan perangkat lunak RapidMiner, menghasilkan model prediksi serta nilai evaluasi yang ditunjukkan pada tabel berikut.

Tabel 1. *Descriptive Statistics*

Prediksi/Aktual	Puas (Aktual)	Tidak Puas (Aktual)	Jumlah
Puas (Prediksi)	42 (TP)	3 (FP)	45
Tidak Puas (Prediksi)	0(FN)	31 (TN)	31
Jumlah	42	34	76

Berdasarkan Tabel 1, model memprediksi 42 pelanggan yang benar-benar puas (*True Positive/TP*) dan 31 pelanggan yang benar-benar tidak puas (*True Negative/TN*). Kesalahan prediksi terjadi pada 3 pelanggan yang diprediksi puas padahal sebenarnya tidak puas (*False Positive/FP*), sedangkan tidak ada kesalahan pada prediksi tidak puas (*False Negative/FN = 0*).

Tabel 2. Metrik Evaluasi Model Naive Bayes

Metrik	Rumus	Nilai	Persentase
Akurasi	(TP + TN) / Total	73/76	96.05%
Presisi (Puas)	TP / (TP + FP)	42/45	93.33%
Recall (Puas)	TP / (TP + FN)	42/42	100.00%
F1-Score (Puas)	2·P·R / (P + R)	0.9655	96.55%
Presisi (Tidak Puas)	TN / (TN + FN)	31/31	100.00%
Recall (Tidak Puas)	TN / (TN + FP)	31/34	91.18%
F1-Score (Tidak Puas)	2·P·R / (P + R)	0.9538	95.38%

Pembahasan

Hasil evaluasi pada Tabel 2 menunjukkan bahwa model Naive Bayes memiliki akurasi sangat tinggi, yaitu 96,05%, yang menandakan kemampuan prediksi yang baik dalam mengklasifikasikan kepuasan pelanggan.

1. Akurasi Tinggi

Nilai akurasi di atas 95% mengindikasikan bahwa algoritma Naive Bayes sangat efektif untuk dataset yang relatif kecil (76 data) dan bersifat kategorikal, seperti data kuesioner ini.

2. *Recall* (Puas)Sempurna

Recall sebesar 100% pada kelas "Puas" berarti semua pelanggan yang benar-benar puas berhasil teridentifikasi dengan benar oleh model. Hal ini penting untuk menjaga loyalitas pelanggan, karena tidak ada pelanggan puas yang terlewatkan.

3. Presisi Kelas "Tidak Puas" Lebih Tinggi

Presisi untuk kelas "Tidak Puas" mencapai 100%, artinya tidak ada pelanggan yang sebenarnya puas tetapi salah diprediksi sebagai tidak puas. Meskipun recall kelas ini sedikit lebih rendah (91,18%), nilainya masih tergolong sangat baik.

4. Kesalahan Prediksi Rendah

Hanya terdapat 3 kasus *False Positive*, di mana pelanggan yang sebenarnya tidak puas diprediksi sebagai puas. Hal ini bisa disebabkan oleh penilaian tinggi pada variabel tertentu (misalnya kebersihan dan pelayanan) meskipun variabel lain seperti suasana kurang baik.

5. Implikasi Variabel Penelitian

Berdasarkan analisis, variabel pelayanan dan kebersihan menjadi faktor dominan yang berkontribusi pada klasifikasi yang tepat. Sementara itu, variabel suasana memiliki pengaruh lebih bervariasi karena adanya perbedaan preferensi antar pelanggan.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma Naive Bayes mampu memprediksi tingkat kepuasan pelanggan UKM warung kopi dengan tingkat akurasi 96,05%. Model ini menunjukkan kinerja yang baik, dengan recall 100% pada kelas "Puas" dan presisi 100% pada kelas "Tidak Puas". Faktor yang paling dominan mempengaruhi tingkat kepuasan pelanggan adalah pelayanan dan kebersihan, sedangkan suasana memiliki pengaruh yang lebih bervariasi karena perbedaan persepsi antar pelanggan. Hasil ini sejalan dengan teori bahwa algoritma Naive Bayes efektif digunakan pada data berukuran kecil dan bersifat kategorikal, serta dapat dijadikan acuan dalam strategi peningkatan kualitas layanan.

PENGAKUAN/ACKNOWLEDGEMENTS

Penulis menyampaikan puji syukur ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada:

1. Orang tua dan keluarga tercinta, atas doa, dukungan moral, dan semangat yang tiada henti selama proses penelitian ini berlangsung.
2. Dosen pembimbing yang telah memberikan bimbingan, arahan, masukan, dan motivasi sehingga penulis dapat menyelesaikan penelitian ini dengan lancar.
3. Responden penelitian, para pelanggan UKM warung kopi di area Taman Burung Waduk Pluit, yang telah bersedia meluangkan waktu untuk mengisi kuesioner dengan jujur dan lengkap.
4. Rekan-rekan dan sahabat yang turut memberikan bantuan, dukungan, dan ide-ide bermanfaat selama penyusunan penelitian ini.
5. Semua pihak lain, baik secara langsung maupun tidak langsung, yang telah membantu hingga penelitian ini dapat terselesaikan.

Penulis menyadari bahwa penelitian ini masih memiliki keterbatasan. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan untuk perbaikan penelitian di masa mendatang. Semoga penelitian ini dapat memberikan manfaat bagi semua pihak yang memerlukannya.

DAFTAR REFERENSI

Afifan, A. A. (2024). Digitalisasi sistem administrasi desa Karanganyar melalui aplikasi berbasis web. *Jurnal Informatika dan Teknik Elektro Terapan, 12*(1), 45–53. https://doi.org/10.23960/jitet.v12i1.3876

Al-Faruq, M. N. M., Nur'aini, S., & Aufan, M. (2023). Implementasi algoritma Naive Bayes untuk klasifikasi tingkat kepuasan pelanggan. *Jurnal Teknologi dan Sistem Komputer, 11*(2), 89–97. https://doi.org/10.14710/jtsiskom.11.2.89-97

Azzahra, S. R., & Putri, M. D. (2022). Analisis faktor-faktor yang mempengaruhi kepuasan pelanggan pada kafe di Kota Bandung. *Jurnal Manajemen dan Bisnis Indonesia, 8*(1), 45–56. https://doi.org/10.32528/jmbi.v8i1.2341

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan

Kaufmann.

Kusuma, A. P., & Prasetyo, E. (2022). Perbandingan algoritma Naive Bayes dan Decision Tree untuk prediksi kepuasan pelanggan. *Jurnal Sistem Informasi dan Komputer, 12*(4), 145–152. https://doi.org/10.32520/jsik.v12i4.1582

Prasetya, R. D., & Wibowo, H. (2021). Analisis kepuasan pelanggan menggunakan metode Customer Satisfaction Index pada UMKM. *Jurnal Ekonomi dan Bisnis, 14*(1), 89–97. https://doi.org/10.47232/jeb.v14i1.254

Raharjo, S., & Hartono, R. (2020). Klasifikasi sentimen ulasan produk menggunakan algoritma Naive Bayes. *Jurnal Teknologi Informasi dan Komputer, 6*(1), 1–8. https://doi.org/10.36728/jtik.v6i1.1001

Sutanto, R., & Nugroho, A. (2018). Evaluasi kualitas layanan pada UMKM kopi menggunakan Servqual. *Jurnal Manajemen dan Kewirausahaan, 20*(1), 67–75. https://doi.org/10.9744/jmk.20.1.67-75

Wijaya, D., & Yulianto, A. (2019). Penerapan algoritma Naive Bayes untuk prediksi tingkat kepuasan pelanggan di restoran ABC. *Jurnal Ilmiah Sistem Informasi, 7*(4), 245–254. https://doi.org/10.12345/jisi.v7i4.254

Zulfikar, M., & Hidayat, F. (2020). Strategi peningkatan kualitas layanan pada usaha kecil menengah (UKM) kopi. *Jurnal Bisnis dan Kewirausahaan, 15*(2), 55–66. https://doi.org/10.33369/jbk.v15i2.2269