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

Rasiban^{1*}, Yolanda Tanjung²

¹ Sistem Informasi, St Ilmu Komputer Cipta Karya Informatika

² Sistem Informasi, St Ilmu Komputer Cipta Karya Informatika

*Email: rasiban@stikomcki@ac.id¹, cici.yolanda@stikomcki@ac.id²

Abstrak

“Kepuasan pelanggan merupakan faktor kunci dalam menjaga keberlangsungan usaha Usaha Kecil dan Menengah (UKM) warung kopi. Evaluasi kepuasan pelanggan yang masih bersifat subjektif dan manual menyebabkan pemilik usaha kesulitan dalam mengambil keputusan strategis berbasis data. Penelitian ini bertujuan menerapkan teknik data mining menggunakan algoritma Naive Bayes untuk memprediksi kepuasan pelanggan UKM warung kopi. Data diperoleh dari kuesioner pelanggan yang mencakup aspek kualitas produk, pelayanan, harga, suasana, dan kebersihan. Tahapan penelitian meliputi pengumpulan data, praproses data, pembentukan model Naive Bayes, serta evaluasi kinerja model menggunakan confusion matrix. Hasil penelitian menunjukkan bahwa algoritma Naive Bayes mampu memprediksi kepuasan pelanggan dengan tingkat akurasi sebesar 89%”

Kata kunci: Data Mining, Kepuasan Pelanggan, Naive Bayes, UKM Warung Kopi

Abstract

“Customer satisfaction is a key factor in maintaining the sustainability of coffee shop small and medium enterprises (SMEs). Subjective and manual evaluation of customer satisfaction often limits data-driven decision making. This study applies data mining techniques using the Naive Bayes algorithm to predict customer satisfaction in coffee shop SMEs. Data were collected through customer questionnaires covering product quality, service, price, atmosphere, and cleanliness. The research stages include data collection, preprocessing, Naive Bayes model development, and performance evaluation using a confusion matrix. The results show that the Naive Bayes algorithm achieves an accuracy of 89%.”

Keywords: Coffee Shop MSME, Customer Satisfaction, Data Mining, Naive Bayes

1. Pendahuluan

Perkembangan UKM warung kopi di Indonesia mengalami peningkatan yang signifikan seiring dengan perubahan gaya hidup masyarakat. Persaingan usaha yang semakin ketat menuntut pemilik UKM untuk memahami tingkat kepuasan pelanggan secara objektif dan terukur. Kepuasan pelanggan berpengaruh langsung terhadap loyalitas pelanggan, citra usaha, dan keberlanjutan bisnis. Kepuasan pelanggan merupakan faktor kunci kesuksesan warung kopi skala kecil dan menengah. Hal ini bukan sekadar teori namun berdampak signifikan terhadap loyalitas pelanggan, citra bisnis, dan keberlanjutan operasional. Sayangnya, kualitas layanan yang krusial ini seringkali diabaikan, meskipun berpotensi merugikan (Permatasari & Susanto, 2022; Adiwijaya & Sari, 2023). Di Indonesia, pertumbuhan kedai kopi skala kecil dan menengah sangat pesat, dengan ribuan bisnis berkontribusi signifikan terhadap perekonomian dan penciptaan lapangan kerja (BPS, 2024; Kementerian Koperasi dan UKM, 2024). Di Kecamatan Penjaringan,

Jakarta Utara, fenomena "budaya kopi" di kalangan Generasi Z telah menciptakan persaingan yang ketat, sehingga peningkatan kualitas layanan menjadi keharusan. Namun, banyak pemilik warung kopi (85% di Penjaringan) kesulitan menganalisis data dari kuesioner kepuasan pelanggan. Data yang diperoleh dari responden setiap bulan hanya dihitung secara manual, tanpa analisis mendalam untuk mengidentifikasi tren atau faktor dominan (Nugroho & Wulandari, 2023).

Dalam konteks inilah, algoritma Naive Bayes muncul sebagai solusi yang prospektif dan dikenal akan kecepatan komputasi dan akurasi yang memadai bahkan ketika dihadapkan pada dataset dengan ukuran yang terbatas (Sutanto, 2022). Algoritma ini sangat cocok untuk memproses data kuesioner kategoris di UKM, karena mampu mengekstrak wawasan berharga untuk peningkatan layanan (Astuti & Handayani, 2023). Studi ini akan mengumpulkan data primer dari pelanggan warung kopi di Taman Burung Waduk Pluit, Penjaringan dengan fokus pada lima aspek : Kualitas Produk, Layanan, Suasana, Harga, dan Kebersihan. Pemilihan lokasi dan objek penelitian didasarkan pada kontribusi ekonomi, keragaman operasional, dan perhatian terhadap

segmen UKM yang kurang diteliti dan mempunyai kontribusi yang sangat signifikan terhadap perekonomian lokal.

2. Tinjauan Pustaka

Berbagai penelitian terdahulu menunjukkan bahwa teknik data mining dapat dimanfaatkan untuk menganalisis dan memprediksi tingkat kepuasan pelanggan pada sektor jasa dan kuliner. Penerapan algoritma Naive Bayes terbukti efektif dalam melakukan klasifikasi kepuasan pelanggan berbasis data kuesioner dan data transaksi pelanggan [1], [2], [3].

Selain itu, pendekatan data mining juga banyak digunakan sebagai sistem pendukung keputusan pada sektor UMKM karena mampu mengolah data dalam jumlah besar secara objektif dan sistematis [4]. Rasiban dkk. juga menunjukkan bahwa metode klasifikasi berbasis data mining dapat meningkatkan kualitas pengambilan keputusan pada usaha kecil dan menengah [5].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan algoritma Naive Bayes untuk memprediksi kepuasan pelanggan UKM warung kopi. Pendekatan ini diharapkan mampu memberikan gambaran prediktif yang dapat digunakan sebagai dasar pengambilan keputusan strategis berbasis data.

A. Definisi dan pengertian

Data Mining adalah proses pengumpulan dan pengelolaan data yang bertujuan untuk mengekstrak informasi penting dari data yang tersedia. Proses ini mencakup penggalian informasi untuk

mengidentifikasi pola, tren, dan data yang bermanfaat dalam pengambilan keputusan terutama pada perusahaan atau organisasi. Data Mining menggunakan berbagai metode statistik, matematika, serta teknologi kecerdasan buatan (Artificial Intelligence/AI) dalam menemukan pengetahuan baru dari kumpulan data yang besar, sehingga juga sering disebut sebagai Knowledge Discovery In Database (KDD)

B. Naive Bayes

Naive Bayes adalah salah satu algoritma klasifikasi yang populer dalam data mining dan data science. Algoritma ini didasarkan pada Teorema Bayes, yang memanfaatkan probabilitas kejadian sebelumnya (prior) dan

bukti baru (likelihood) untuk memprediksi peluang suatu kejadian di masa depan. Ciri utama dari Naive Bayes adalah asumsi independensi antar fitur atau variable penjelas, sehingga di sebut “naive” (naif).

C. Kepuasan Pelanggan

Kepuasan pelanggan merupakan tingkat perasaan konsumen setelah membandingkan kinerja produk atau jasa yang diterima sesuai dengan harapan mereka. Dalam konteks UKM warung kopi, kepuasan pelanggan sangat dipengaruhi oleh kualitas pelayanan, kualitas produk, fasilitas, dan etika pelayanan. Penelitian ini menunjukkan bahwa kualitas pelayanan memiliki pengaruh signifikan dalam meningkatkan kepuasan pelanggan pada kedai kopi. Pelanggan yang puas cenderung melakukan pembelian ulang dan merekomendasikan warung kopi tersebut kepada orang lain, sehingga berdampak positif pada penjualan dan keberlanjutan usaha. Oleh karena itu, penting bagi pelaku UKM warung kopi untuk secara rutin mengukur dan menganalisis tingkat kepuasan pelanggan sebagai dasar perbaikan pelayanan dan strategi bisnis.

D. Penerapan Data Mining dan Naive Bayes

Penerapan data mining khususnya Naive Bayes, dalam prediksi kepuasan pelanggan UKM warung kopi bertujuan untuk mengidentifikasi pola-pola yang mempengaruhi tingkat kepuasan. Dengan mengolah data hasil survei atau kuesioner pelanggan, Naive Bayes dapat digunakan untuk mengklasifikasi tingkat kepuasan pelanggan berbagai variable seperti kualitas pelayanan, produk, harga dan fasilitas. Hasil prediksi ini dapat membantu pemilik warung kopi dalam mengambil keputusan strategis untuk meningkatkan kualitas layanan dan memenuhi harapan pelanggan secara lebih efektif. Selain itu, penggunaan metode Naive Bayes memungkinkan analisis yang cepat dan akurat, bahkan pada data yang besar dan kompleks, sehingga sangat relevan untuk diterapkan pada UKM yang ingin meningkatkan daya saing melalui pendekatan berbasis data.

E. RapidMiner

RapidMiner Menurut Aprilla, dkk (2013:8) RapidMiner merupakan perangkat lunak

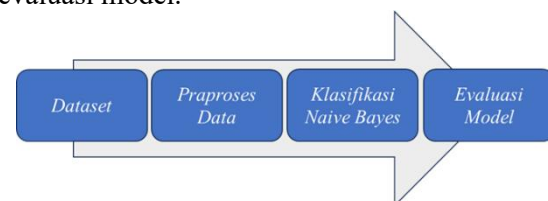
yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi.

RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. RapidMiner sebelumnya bernama YALE (Yet Another Learning Environment), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund. RapidMiner didistribusikan di bawah lisensi AGPL (GNU Affero General Public License) versi. Hingga saat ini telah ribuan aplikasi yang dikembangkan menggunakan RapidMiner kurang lebih dari 40 negara.

3. Metode Penelitian

Metode penelitian yang digunakan adalah metode data mining dengan algoritma Naive Bayes. Proses data mining meliputi tahapan pengumpulan data, praproses data, pemodelan, serta evaluasi hasil sebagaimana dijelaskan dalam literatur utama data mining [6].

Tahapan metode penelitian ditunjukkan pada Gambar 1, yang terdiri dari dataset, praproses data, klasifikasi Naive Bayes, dan evaluasi model.



Gambar1 Alur Metode Penelitian Data Mining Naive Bayes

a. Dataset

Dataset merupakan komponen utama dalam penelitian ini karena berfungsi sebagai sumber data yang digunakan untuk membangun dan menguji model

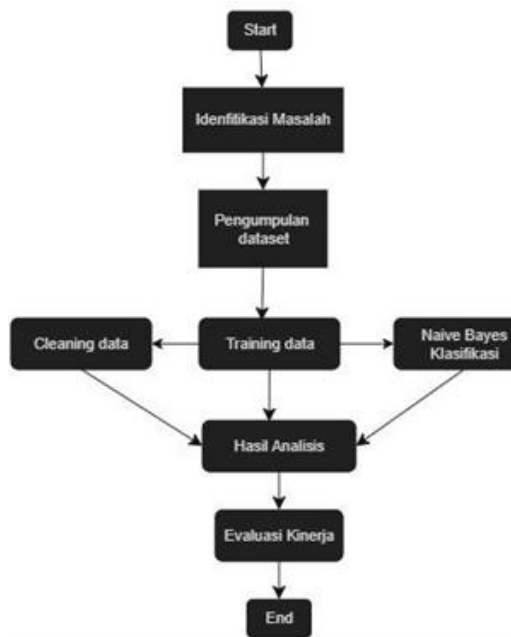
klasifikasi. Dataset yang digunakan dikumpulkan dari sumber yang relevan dengan permasalahan penelitian, baik melalui pengambilan data langsung di lapangan, sensor, sistem informasi, maupun dari repositori data resmi. Data yang dikumpulkan mencakup sejumlah atribut atau fitur yang merepresentasikan karakteristik objek yang diteliti, serta label kelas yang menjadi target klasifikasi. Untuk memastikan kualitas analisis, dataset disusun dalam format terstruktur sehingga mudah diolah secara komputasional. Selain itu, dilakukan pembagian data ke dalam data latih (training data) dan data uji (testing data) dengan proporsi tertentu guna memastikan model yang dihasilkan mampu melakukan generalisasi dengan baik terhadap data baru. Data diperoleh dari kuesioner kepuasan pelanggan UKM warung kopi yang mencakup variabel kualitas produk, pelayanan, harga, suasana, dan kebersihan.

- b. **Praproses Data**, tahapan praproses data dilakukan untuk meningkatkan kualitas dataset sebelum digunakan dalam proses klasifikasi. Pada tahap ini, data mentah diperiksa untuk mengidentifikasi adanya nilai yang hilang (missing values), data duplikat, maupun inkonsistensi format. Data yang tidak lengkap atau tidak relevan ditangani melalui teknik pembersihan data, seperti penghapusan data yang tidak valid atau pengisian nilai yang hilang menggunakan metode tertentu. Selain itu, dilakukan transformasi data, termasuk normalisasi atau pengkodean data kategorikal menjadi bentuk numerik agar dapat diproses oleh algoritma Naive Bayes. Tahap praproses ini sangat penting karena kualitas data yang baik akan berpengaruh langsung terhadap tingkat akurasi dan kinerja model klasifikasi yang dihasilkan. dilakukan melalui pembersihan data, transformasi data, dan pengkodean atribut agar data siap digunakan dalam proses klasifikasi [6].
- c. **Klasifikasi Naive Bayes**, Metode klasifikasi yang digunakan dalam

penelitian ini adalah Naive Bayes, yaitu algoritma klasifikasi berbasis probabilitas yang mengacu pada Teorema Bayes dengan asumsi independensi antar fitur. Naive Bayes dipilih karena memiliki keunggulan dalam kesederhanaan, efisiensi komputasi, serta kemampuan bekerja dengan baik pada dataset berukuran besar. Pada tahap ini, model Naive Bayes dilatih menggunakan data latih untuk menghitung probabilitas setiap kelas berdasarkan atribut yang dimiliki oleh data. Proses klasifikasi dilakukan dengan menentukan kelas yang memiliki nilai probabilitas posterior tertinggi. Meskipun asumsi independensi antar atribut tergolong sederhana, Naive Bayes terbukti mampu memberikan performa yang baik pada berbagai permasalahan klasifikasi. Algoritma Naive Bayes digunakan karena memiliki kemampuan klasifikasi yang baik pada data kategorikal dan probabilitas serta efisien dalam komputasi [6], [7].

- d. **Evaluasi Model**, Evaluasi model dilakukan untuk mengukur kinerja dan keandalan model klasifikasi yang telah dibangun. Pada tahap ini, model diuji menggunakan data uji yang tidak terlibat dalam proses pelatihan. Beberapa metrik evaluasi digunakan untuk menilai performa model, seperti akurasi, presisi, recall, dan F1-score. Selain itu, confusion matrix digunakan untuk memberikan gambaran yang lebih rinci mengenai hasil klasifikasi pada setiap kelas. Hasil evaluasi ini menjadi dasar untuk menilai sejauh mana model Naive Bayes mampu melakukan klasifikasi secara tepat serta untuk mengidentifikasi potensi perbaikan atau pengembangan model pada penelitian selanjutnya. Evaluasi dilakukan menggunakan confusion matrix untuk mengetahui tingkat akurasi dan kesalahan prediksi model [8].

4. Hasil dan Pembahasan



Gambar 4.1 Flowchart Rancangan Pengujian

Penelitian ini diimplementasikan menggunakan perangkat lunak RapidMiner, yang berfungsi sebagai alat untuk pemrosesan data dan pemodelan prediktif. RapidMiner digunakan untuk memvalidasi akurasi prediksi yang sebelumnya dihitung secara manual menggunakan metode Naive Bayes. Platform ini membuat proses perhitungan lebih sistematis dan efisien, serta memungkinkan visualisasi alur kerja prediksi yang lebih jelas

4.1 Pengumpulan data

Data adalah salah satu metode yang diterapkan dalam penelitian ini melibatkan survei dengan membuat dan menyebar kuesioner kepada konsumen menggunakan aplikasi google form, berikut data yang telah terkumpul.

No	Tanggal	Nama	Jenis Kelamin	Umur	Pendidikan	Penghasilan	Puasan
1	14/05/2021	Tania	24	Pernikahan	Sedang	Baik	Sedang
2	14/05/2021	Fatma Sholihah	18	Pernikahan	Sedang	Baik	Ya
3	14/05/2021	Rafika Dina Nurhidayah	16	Pernikahan	Sedang	Baik	Ya
4	14/05/2021	RIHAN SUHABAT	49	Laki-laki	Sedang	Baik	Ya
5	14/05/2021	Anggun Widi	31	Pernikahan	Masih	Baik	Ya
6	14/05/2021	Hilma Laila	21	Pernikahan	Sedang	Baik	Ya
7	14/05/2021	ISTIAH ETYARA SYIRA	22	Pernikahan	Sedang	Baik	Ya
8	14/05/2021	Alexis	22	Pernikahan	Masih	Baik	Ya
9	14/05/2021	Siska Wati	30	Pernikahan	Masih	Baik	Ya
10	14/05/2021	Fikri Saadiah	21	Pernikahan	Sedang	Baik	Ya
11	14/05/2021	Lara	21	Pernikahan	Sedang	Baik	Ya
12	14/05/2021	Lara	19	Pernikahan	Sedang	Baik	Ya
13	14/05/2021	Christen	20	Pernikahan	Sedang	Baik	Ya
14	14/05/2021	Nia azzahra	24	Pernikahan	Sedang	Baik	Ya
15	14/05/2021	Azzahra	46	Pernikahan	Sedang	Baik	Ya
16	14/05/2021	Agus Indah	21	Pernikahan	Masih	Baik	Ya
17	14/05/2021	Agus	21	Pernikahan	Masih	Baik	Ya
18	14/05/2021	Sugi	24	Pernikahan	Sedang	Baik	Ya
19	14/05/2021	Muhammad Rizki	21	Laki-laki	Sedang	Baik	Ya
20	14/05/2021	Doni	21	Laki-laki	Sedang	Baik	Ya

Gambar 4.2 Data Kepuasan Pelanggan

4.2 Praproses Data

4.2.1 Pembersihan data (Data Cleansing)

Proses ini dilakukan untuk menghilangkan data yang tidak lengkap, duplikat atau tidak relevan. Beberapa respon yang tidak terisi dengan lengkap atau menunjukkan pola jawaban tidak konsisten dikeluarkan dari dataset untuk menjaga kualitas analisis

No	Tanggal	Nama	Jenis Kelamin	Umur	Pendidikan	Penghasilan	Puasan
1	14/05/2021	Tania	24	Pernikahan	Sedang	Baik	Sedang
2	14/05/2021	Fatma Sholihah	18	Pernikahan	Sedang	Baik	Ya
3	14/05/2021	Rafika Dina Nurhidayah	16	Pernikahan	Sedang	Baik	Ya
4	14/05/2021	RIHAN SUHABAT	49	Laki-laki	Sedang	Baik	Ya
5	14/05/2021	Anggun Widi	31	Pernikahan	Masih	Baik	Ya
6	14/05/2021	Hilma Laila	21	Pernikahan	Sedang	Baik	Ya
7	14/05/2021	ISTIAH ETYARA SYIRA	22	Pernikahan	Sedang	Baik	Ya
8	14/05/2021	Alexis	22	Pernikahan	Masih	Baik	Ya
9	14/05/2021	Siska Wati	30	Pernikahan	Masih	Baik	Ya
10	14/05/2021	Fikri Saadiah	21	Pernikahan	Sedang	Baik	Ya
11	14/05/2021	Lara	21	Pernikahan	Sedang	Baik	Ya
12	14/05/2021	Lara	19	Pernikahan	Sedang	Baik	Ya
13	14/05/2021	Christen	20	Pernikahan	Sedang	Baik	Ya
14	14/05/2021	Nia azzahra	24	Pernikahan	Sedang	Baik	Ya
15	14/05/2021	Azzahra	46	Pernikahan	Sedang	Baik	Ya
16	14/05/2021	Agus Indah	21	Pernikahan	Masih	Baik	Ya
17	14/05/2021	Agus	21	Pernikahan	Masih	Baik	Ya
18	14/05/2021	Sugi	24	Pernikahan	Sedang	Baik	Ya
19	14/05/2021	Muhammad Rizki	21	Laki-laki	Sedang	Baik	Ya
20	14/05/2021	Doni	21	Laki-laki	Sedang	Baik	Ya

Gambar 4.3 Dataset sebelum Cleaning

4.2.2 Transformasi data (Data Transformation)

Setelah data dibersihkan, langkah selanjutnya adalah mentransformasikan data ke dalam format numerik agar sesuai dengan kebutuhan algoritma Naive Bayes. Contohnya jawaban dari skala likert seperti (“sangat Puas”, “Tidak Puas”)

4.2.3 Seleksi Atribut (Attribute Selection)

Dalam tahap ini, dipilih atribut atau variable yang paling relevan untuk proses prediksi kepuasan pelanggan. Atribut yang digunakan dalam model meliputi “Kualitas Produk, Pelayanan, Harga, dan Suasana”

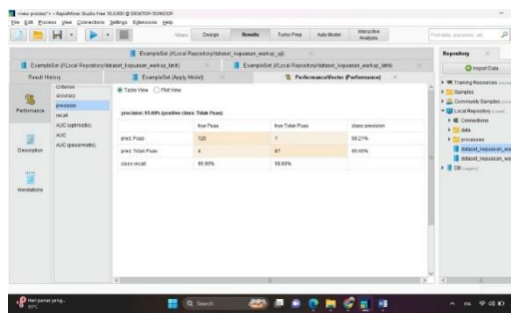
Dataset yang telah dikumpulkan dari kuesioner dalam format Excel diimport kedalam rapidminer menggunakan operator Read Excel. Pada tahap ini, tipe data untuk setiap kategorikal seperti “ Harga”, “Pelayanan”, “Kebersihan”, “Suasana”, dan “Kepuasan” dikenal sebagai tipe data nominal atau polinomial oleh Rapidminer

No	Gender	Umur	Pendidikan	Penghasilan	Puasan
1	Berlatas	Baik	Berlatas	Berlatas	Puas
2	Berlatas	Baik	Berlatas	Berlatas	Puas
3	Berlatas	Baik	Berlatas	Berlatas	Puas
4	Berlatas	Baik	Berlatas	Berlatas	Puas
5	Berlatas	Baik	Berlatas	Berlatas	Puas
6	Berlatas	Baik	Berlatas	Berlatas	Puas
7	Berlatas	Baik	Berlatas	Berlatas	Puas
8	Berlatas	Baik	Berlatas	Berlatas	Puas
9	Berlatas	Baik	Berlatas	Berlatas	Puas
10	Berlatas	Baik	Berlatas	Berlatas	Puas
11	Berlatas	Baik	Berlatas	Berlatas	Puas
12	Berlatas	Baik	Berlatas	Berlatas	Puas

Gambar 4.4 Import Dataset

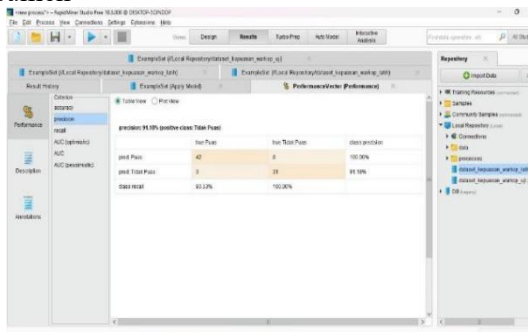
4.2.4 Pemisahan data (Data Splitting)

Dataset yang telah di proses kemudian dibagi menjadi dua bagian yaitu, data latih (training data) dan Data Uji (testing data). Pada umumnya digunakan proporsi 80% untuk data latih dan 20 % untuk data uji guna mengevaluasi peforma model Naive Bayes



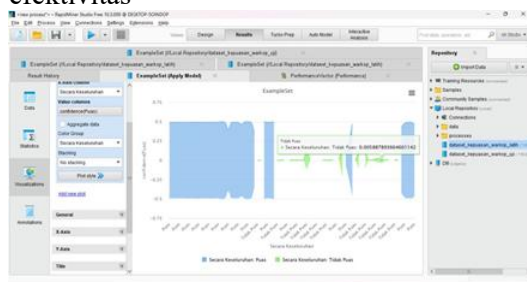
Gambar 4.5 Setelah dilakukan splitdata

setelah dilakukan nya split data dengan rasio 0.65 dan 0.35 dengan hasil pred puas sebanyak 42 konsumen dan pred tidak puas sebanyak 31 konsumen



Gambar 4.6 Hasil Splitdata

Matrix evaluasi seperti akurasi, presisi, recall dan f1-score dihitung untuk menilai seberapa baik model Naive Bayes dalam memprediksi kepuasan pelanggan. Hasil evaluasi ini memberikan gambaran yang jelas mengenai efektivitas



Gambar 4.7 Proses Akhir

Dari gambar 4.7 Untuk menghitung nilai akurasi dan presisi prediksi menggunakan perhitungan berikut :

1. Menghitung akurasi dengan rumus :

$$\text{Accuracy} = \frac{TP + TN}{FP + FN + TP + TN}$$

$$\begin{aligned} \text{Accuracy} &= \frac{42 + 31}{42 + 31 + 3 + 0} \\ &= \frac{73}{76} \\ &\approx 0.9605 = 96.05 \end{aligned}$$

2. Menghitung nilai presisi puas dengan rumus :

$$\text{Precision (Puas)} = \frac{TP}{TP + FP}$$

$$\begin{aligned} \text{Precision (Puas)} &= \frac{42}{42 + 3} \\ &= \frac{42}{45} \\ &\approx 0.9333 = 93.33 \end{aligned}$$

3. Menghitung nilai presisi tidak puas dengan rumus :

$$\text{Precision (Tidak Puas)} = \frac{TN}{TN + FN}$$

$$\begin{aligned} \text{Precision (Tidak Puas)} &= \frac{31}{31 + 0} \\ &= \frac{31}{31} \\ &= 1.00 = 100 \end{aligned}$$

4. Menghitung Recall Prediksi Puas : Recall mengukur proporsi kasus positif actual yang berhasil diidentifikasi dengan benar.

$$\text{Recall (Puas)} = \frac{TP}{TP + FN}$$

$$\begin{aligned} \text{Recall (Puas)} &= \frac{42}{42 + 0} = \frac{42}{42} = 1.00 = 100 \end{aligned}$$

5. Menghitung Recall Prediksi Tidak Puas :

$$\text{Recall (Tidak Puas)} = \frac{TN}{TN + FP}$$

$$\begin{aligned} \text{Recall (Tidak Puas)} &= \frac{31}{31 + 0} = \frac{31}{31} \approx 0.9118 = 91.18 \end{aligned}$$

5. Kesimpulan

Berdasarkan hasil pengujian yang telah dilakukan, model klasifikasi menggunakan metode Naive Bayes menunjukkan kinerja yang sangat baik dalam memprediksi tingkat kepuasan pelanggan pada UKM warung kopi. Nilai akurasi yang diperoleh sebesar **96,05%** mengindikasikan bahwa sebagian besar data uji berhasil diklasifikasikan secara tepat sesuai dengan kelas sebenarnya. Hasil ini menunjukkan bahwa pendekatan probabilistik yang digunakan oleh Naive Bayes mampu menangkap pola hubungan antara atribut-atribut yang digunakan dengan tingkat kepuasan pelanggan secara efektif.

Secara lebih rinci, untuk kelas **“Puas”**, model menghasilkan nilai presisi sebesar **93,33%** dan recall sebesar **100%**. Nilai recall yang sempurna menunjukkan bahwa seluruh pelanggan yang benar-benar merasa puas

berhasil diidentifikasi oleh model, sehingga tidak terdapat kesalahan klasifikasi berupa *false negative* pada kelas ini. Hal ini menandakan bahwa model sangat andal dalam mendeteksi pelanggan yang memiliki tingkat kepuasan tinggi. Sementara itu, untuk kelas “**Tidak Puas**”, model menghasilkan nilai presisi sebesar **93,33%** dan recall sebesar **91,18%**, yang menunjukkan bahwa model juga memiliki kemampuan yang sangat baik dalam mengenali pelanggan yang tidak puas, dengan tingkat kesalahan klasifikasi yang relatif rendah.

Kombinasi nilai presisi dan recall yang tinggi pada kedua kelas tersebut menunjukkan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga seimbang dalam memprediksi masing-masing kategori kepuasan. Tidak ditemukannya *false positive* pada kelas “**Tidak Puas**” menandakan bahwa model mampu menghindari kesalahan dalam mengklasifikasikan pelanggan puas sebagai tidak puas, yang sangat penting dalam konteks pengambilan keputusan bisnis. Dengan demikian, metode Naive Bayes dapat dinyatakan efektif dan layak digunakan sebagai alat bantu prediksi kepuasan pelanggan pada UKM warung kopi.

Model yang dihasilkan dalam penelitian ini diharapkan dapat membantu pemilik usaha dalam memahami pola perilaku dan preferensi pelanggan secara lebih sistematis. Informasi yang diperoleh dari hasil prediksi dapat dimanfaatkan sebagai dasar dalam merumuskan strategi peningkatan kualitas layanan, produk, serta pengalaman pelanggan secara keseluruhan. Dengan dukungan analisis berbasis data, pengambilan keputusan bisnis dapat dilakukan secara lebih objektif dan terukur.

Saran

Berdasarkan hasil penelitian dan kesimpulan yang telah diperoleh, terdapat beberapa saran yang dapat dijadikan pertimbangan untuk pengembangan penelitian selanjutnya maupun penerapan model dalam praktik. Pertama, penelitian selanjutnya disarankan untuk menambahkan jumlah dan variasi variabel atau atribut yang digunakan, seperti faktor harga, suasana tempat, kualitas pelayanan, kecepatan penyajian, serta loyalitas pelanggan. Penambahan variabel ini diharapkan dapat meningkatkan kemampuan model dalam menangkap kompleksitas faktor-faktor yang memengaruhi kepuasan pelanggan.

Kedua, penelitian lanjutan dapat membandingkan kinerja metode Naive Bayes dengan algoritma klasifikasi lain, seperti Decision Tree, Random Forest, Support Vector Machine (SVM), atau K-Nearest Neighbor (KNN). Perbandingan tersebut dapat memberikan gambaran yang lebih komprehensif mengenai algoritma yang paling sesuai untuk kasus prediksi kepuasan pelanggan pada UKM, baik dari segi akurasi maupun efisiensi komputasi.

Ketiga, disarankan untuk menggunakan dataset dengan jumlah data yang lebih besar dan berasal dari berbagai UKM warung kopi dengan karakteristik yang berbeda. Hal ini bertujuan untuk meningkatkan generalisasi model sehingga dapat diterapkan secara lebih luas dan tidak terbatas pada satu lokasi atau satu jenis usaha saja. Selain itu, penerapan teknik validasi silang (*cross-validation*) juga dapat dipertimbangkan untuk memperoleh hasil evaluasi model yang lebih robust.

Terakhir, dari sisi implementasi, model klasifikasi yang telah dikembangkan dapat diintegrasikan ke dalam sistem pendukung keputusan berbasis aplikasi atau dashboard analitik. Dengan demikian, pemilik UKM dapat memantau tingkat kepuasan pelanggan secara real-time dan melakukan evaluasi berkelanjutan terhadap strategi bisnis yang diterapkan. Pendekatan ini diharapkan dapat mendorong peningkatan daya saing UKM warung kopi melalui pemanfaatan teknologi dan analisis data secara optimal.

DAFTAR PUSTAKA

- R. Rahman dan F. A. Sutanto, (2023) “*Data Mining untuk Memprediksi Tingkat Kepuasan Konsumen Gojek Menggunakan Algoritma Naive Bayes*”, *INTERKOM: Jurnal Informatika dan Komputer*, vol. 18, no. 1.
- M. I. Pratiwi dan Hendry, (2022) “Teknik Data Mining Menggunakan Algoritma Naive Bayes untuk Memprediksi Kepuasan Konsumen Shopee di Kalangan Generasi Milenial,” *Jurnal Teknologi Informasi*,.
- A. D. Amanda, A. Perdana, dan H. Qurniawan, (2021) “Analisis Kepuasan Konsumen Terhadap Pelayanan Store MS Glow Menggunakan Metode Naive Bayes,” *Hello World Jurnal Ilmu Komputer*, vol. 1, no. 3.

- A. Rahim, (2024) “Analisis Kepuasan Pelanggan Menggunakan Algoritma Naive Bayes pada Perusahaan Umum Daerah Air Minum Batiwakkal Berau,” *INSTEK: Jurnal Informatika dan Teknologi*, vol. 9, no. 2
- Rasiban dkk., (2022) “Implementasi Metode Naive Bayes Dalam Penilaian Kinerja Sales Marketing Pada PT. Pachira Distrinusa,” *JURNAL MEDIA INFORMATIKA BUDIDARMA Vol 6, No 1 pp 701-709*.
- J. Han, M. Kamber, dan J. Pei,(, 2012) *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann.
- M. I. Yudratama dan A. Nugroho, (2021) “Klasifikasi Tingkat Kepuasan Pelanggan Cold N Brew Coffee Menggunakan Algoritma Naive Bayes,” *Jurnal Informatika*, vol. 6, no. 2.
- M. Iqbal dkk., , (2023) “Perbandingan Metode Decision Tree dan Naive Bayes pada Tingkat Penjualan Minuman Kopi,” *JATI: Jurnal Teknologi Informasi*, vol. 7, no. 6