

**ALGORITMA NAIVE BAYES DAN KNN DALAM  
KLASIFIKASI KEPUASAN PELANGGAN**



# ALGORITMA NAIVE BAYES DAN KNN DALAM KLASIFIKASI KEPUASAN PELANGGAN

Windi Duwi Cahyani  
Sudi Suryadi, S.Kom., M.Kom  
Syaiful Zuhri Harahap, S.Kom., M.Kom  
Ibnu Rasyid Munthe, S.T., M.Kom



# ALGORITMA NAIVE BAYES DAN KNN DALAM KLASIFIKASI KEPUASAN PELANGGAN

@Windi Duwi Cahyani | Sudi Suryadi, S.Kom., M.Kom | Syaiful Zuhri Harahap,  
S.Kom., M.Kom | Ibnu Rasyid Munthe, S.T., M.Kom

Pemerhati Aksara — Tim Penerbit Pustaka Riyadz  
Perancang Sampul — Tim Penerbit Pustaka Riyadz  
Penata Letak— Tim Penerbit Pustaka Riyadz

Cetakan Pertama, Maret 2026  
15,5 × 23 cm  
ISBN — 978-634-7078-70-4

Diterbitkan oleh:



**CV. Pustaka Riyadz**

Desa Sawahan Jaya, Semurup,  
Kec. Air Hangat, Kerinci - Jambi  
pustakariyadz@gmail.com  
Instagram @pustakariyadz  
HP 0852-6988-7472  
Anggota IKAPI : No. 08/JBI/2023

Didistribusi Oleh:



**Pendekar Publisher**

Jl. Bariang Indah 3G, Anduring,  
Kec. Kuranji, Padang - Sumbar  
pendekarpublisher@gmail.com  
Instagram @pendekarpublisher  
HP 0897-1774-592

*Hak cipta dilindungi undang-undang;  
Dilarang mengutip atau memperbanyak sebagian atau seluruh isi buku ini  
tanpa seizin tertulis dari Penerbit*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas rahmat dan karunia-Nya buku yang berjudul "Algoritma Naive Bayes dan KNN dalam Klasifikasi Kepuasan Pelanggan" ini dapat diselesaikan dengan baik. Buku ini disusun sebagai upaya untuk memberikan pemahaman mengenai penerapan metode klasifikasi dalam bidang data mining dan machine learning, khususnya dalam menganalisis tingkat kepuasan pelanggan menggunakan algoritma Naive Bayes dan K-Nearest Neighbor (KNN).

Dalam era digital saat ini, data menjadi salah satu aset penting bagi organisasi maupun perusahaan. Salah satu pemanfaatan data yang sangat penting adalah untuk memahami tingkat kepuasan pelanggan, sehingga perusahaan dapat meningkatkan kualitas layanan dan pengambilan keputusan yang lebih tepat. Melalui buku ini, penulis berusaha menjelaskan konsep dasar klasifikasi data, prinsip kerja algoritma Naive Bayes dan KNN, serta implementasinya dalam studi kasus klasifikasi kepuasan pelanggan secara sistematis dan mudah dipahami.

Buku ini diharapkan dapat menjadi referensi bagi mahasiswa, peneliti, maupun praktisi yang tertarik dalam bidang data mining, machine learning, dan analisis kepuasan pelanggan. Penulis menyadari bahwa buku ini masih memiliki berbagai keterbatasan, baik dari segi penyajian maupun kedalaman materi. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan untuk penyempurnaan di masa yang akan datang.

Akhir kata, penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dalam penyusunan buku ini. Semoga buku ini bermanfaat bagi pengembangan ilmu pengetahuan, khususnya dalam penerapan algoritma klasifikasi pada analisis kepuasan pelanggan.

# DAFTAR ISI

<b>KATA PENGANTAR .....</b>	<b>v</b>
<b>DAFTAR ISI .....</b>	<b>vi</b>
<b>BAB 1 .....</b>	<b>3</b>
<b>KONSEP DATA MINING DAN KLASIFIKASI .....</b>	<b>3</b>
1. Knowledge Discovery in Database (KDD) .....	3
2. Data Mining .....	6
3. Model Klasifikasi .....	8
<b>BAB 2 .....</b>	<b>13</b>
<b>ALGORITMA NAIVE BAYES .....</b>	<b>13</b>
1. Pengertian Naive Bayes .....	13
2. Konsep Dasar Metode Naive Bayes .....	15
3. Kelebihan Metode Naive Bayes .....	16
4. Kekurangan Metode Naive Bayes .....	18
5. Pengolahan Data Metode Naive Bayes .....	20
<b>BAB 3 .....</b>	<b>25</b>
<b>ALGORITMA K-NEAREST NEIGHBOR (KNN) .....</b>	<b>25</b>
1. Pengertian K-Nearest Neighbor .....	25
2. Konsep Dasar Metode KNN .....	27
3. Kelebihan Metode KNN .....	29
4. Kekurangan Metode KNN .....	31
5. Pengolahan Data Metode KNN .....	32
<b>BAB 4 .....</b>	<b>37</b>
<b>TOOLS PENGUJIAN ALGORITMA NAIVE BAYES DAN KNN .....</b>	<b>37</b>
1. RapidMiner .....	37
2. Microsoft Excel .....	38
<b>BAB 5 .....</b>	<b>43</b>
<b>IMPLEMENTASI ALGORITMA NAIVE BAYES DAN KNN PADA KLASIFIKASI KEPUASAN PELANGGAN .....</b>	<b>43</b>
1. Latar Belakang .....	43
2. Gambaran Umum Objek Penelitian .....	46
3. Hasil dan Pembahasan .....	57
4. Penutup .....	66
<b>DAFTAR PUSTAKA .....</b>	<b>71</b>

# BAB 1





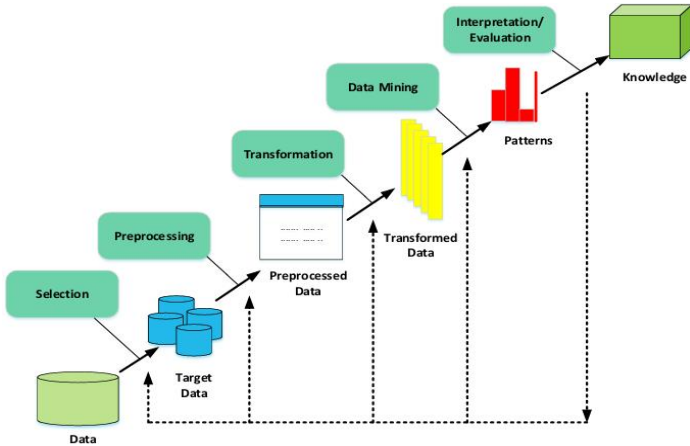
## **BAB 1**

### **KONSEP DATA MINING DAN KLASIFIKASI**

#### **1. Knowledge Discovery in Database (KDD)**

Knowledge Discovery in Database (KDD) merupakan suatu pendekatan sistematis untuk mengambil informasi yang penting, tersirat, dan belum diketahui dari sekumpulan data (Kharis et al., 2023). KDD dimulai dari proses seleksi data, pembersihan data untuk menghilangkan noise atau inkonsistensi, transformasi data ke dalam bentuk yang lebih sesuai untuk dianalisis, hingga penerapan metode Data mining untuk menemukan pola, hubungan, atau informasi bernilai. Tahapan terakhir adalah interpretasi dan evaluasi, di mana pola yang ditemukan dinilai kembali agar dapat dipastikan relevansi dan kegunaannya sebelum diubah menjadi pengetahuan yang siap digunakan. KDD tidak hanya fokus pada teknik algoritmik, tetapi juga memastikan kualitas data, validitas pola, serta kemampuan informasi yang dihasilkan untuk mendukung pemahaman dan pengambilan keputusan yang lebih baik.

Penerapan KDD memberikan dasar metodologis yang kuat untuk menggali pola kepuasan pelanggan dari data yang dimiliki Mie Gacoan Rantauprapat. Proses KDD adalah proses pengumpulan, penggunaan data untuk menentukan aturan atau pola tertentu dalam jumlah data yang sangat besar. Tahapan KDD membantu memastikan bahwa data yang dianalisis benar-benar mencerminkan perilaku pelanggan, sehingga pola klasifikasi yang dihasilkan memiliki akurasi yang optimal. Selain itu, proses evaluasi dalam KDD memungkinkan hasil prediksi dari kedua metode dibandingkan secara objektif, sehingga dapat diperoleh pengetahuan baru mengenai metode yang memberikan performa paling baik untuk memetakan tingkat kepuasan pelanggan.



**Gambar 1.** Proses Knowledge Discovery in Database (KDD)  
(Purwati et al., 2023)

Pada gambar diatas merupakan proses yang dilakukan KDD, proses diatas memiliki fungsi dan maksudnya masing masing. Untuk penjelasan dari setiap tahapan diatas akan dijabarkan dibawah ini.

### 1. Seleksi Data (Data Selection)

Seleksi data merupakan proses memilih data yang akan digunakan. Data yang dipilih kemudian dibagi menjadi data pelatihan dan data pengujian (Bustomi et al., 2023).

### 2. Pembersihan Data (Data Cleaning)

Pembersihan data dilakukan untuk menghilangkan data yang tidak lengkap, tidak konsisten, atau mengandung kesalahan. Kegiatan ini meliputi penanganan data kosong, data ganda, serta kesalahan input. Data yang telah dibersihkan akan menghasilkan analisis yang lebih akurat.

### 3. Transformasi Data (Data Transformation)

Transformasi data bertujuan untuk mengubah data ke dalam format yang sesuai untuk proses data mining. Transformasi dapat berupa normalisasi, pengelompokan nilai, atau perubahan data

kategorikal menjadi numerik. Proses ini membantu meningkatkan kinerja algoritma klasifikasi.

#### **4. Data Mining (Data Mining)**

Data mining merupakan proses untuk menemukan pola atau informasi penting dari sekumpulan data. Pada penelitian ini digunakan metode Naive Bayes dan K-Nearest Neighbor (KNN) untuk mengklasifikasikan tingkat kepuasan pelanggan. Hasil dari proses ini berupa model klasifikasi yang dapat digunakan untuk analisis dan prediksi.

#### **5. Evaluasi dan Interpretasi (Evaluation and Interpretation)**

Evaluasi dan interpretasi dilakukan untuk menilai hasil dari proses data mining. Penilaian dilakukan dengan menggunakan ukuran kinerja seperti akurasi untuk mengetahui keefektifan metode yang digunakan. Hasil evaluasi kemudian diinterpretasikan agar dapat dijadikan dasar pengambilan keputusan.

Berdasarkan tahapan KDD yang telah dijelaskan, setiap proses memiliki peranan penting dalam menghasilkan pengetahuan yang berkualitas dan akurat. Dimulai dari seleksi data hingga evaluasi dan interpretasi, seluruh tahapan saling berkaitan dan tidak dapat dipisahkan satu sama lain. Data yang dipilih dan dibersihkan dengan baik akan sangat memengaruhi keberhasilan proses data mining, karena kualitas data menjadi faktor utama dalam menghasilkan model klasifikasi yang optimal. Transformasi data juga berperan penting dalam menyesuaikan format data agar dapat diproses dengan baik oleh algoritma Naive Bayes dan K-Nearest Neighbor (KNN), sehingga proses klasifikasi dapat berjalan secara efektif dan efisien.

KDD adalah proses yang tidak sepele yang digunakan dalam pengenalan pola atau hubungan yang valid, baru, memiliki potensi manfaat, dan pada akhirnya dapat dipahami dalam kumpulan data untuk pengambilan keputusan penting (Hermawan & Prianggono, 2023). Melalui proses data mining, model klasifikasi

yang dihasilkan mampu memberikan gambaran tingkat kepuasan pelanggan berdasarkan data yang telah diolah. Selanjutnya, evaluasi dan interpretasi hasil memberikan pemahaman yang lebih mendalam mengenai kinerja masing-masing metode yang digunakan. Dengan demikian, hasil dari keseluruhan proses KDD diharapkan dapat menjadi dasar yang kuat dalam pengambilan keputusan serta memberikan rekomendasi yang bermanfaat bagi pihak manajemen dalam upaya meningkatkan kualitas pelayanan dan kepuasan pelanggan.

## **2. Data Mining**

Data mining adalah proses mengklasifikasikan data dengan mengelompokkan data sesuai dengan kelasnya masing-masing (M. Sari et al., 2023). Proses ini mencakup tahapan seperti pembersihan data, transformasi, pemilihan metode, hingga interpretasi hasil agar dapat menghasilkan pengetahuan baru yang bermakna. Data mining proses semi-otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi pengetahuan potensial dan bermanfaat yang tersimpan dalam basis data besar (Abas et al., 2023). Dengan kemampuannya mengolah data besar menjadi informasi yang lebih mudah dipahami, Data mining menjadi salah satu fondasi utama dalam pengambilan keputusan berbasis data di berbagai bidang mulai dari bisnis, kesehatan, pendidikan hingga sistem informasi modern.

Data mining adalah suatu teknik analisis yang memanfaatkan teknologi pemrosesan data dan metodologi statistik untuk menemukan pola dan hubungan dalam data (Kurniadi Hermawan et al., 2023). Penggunaan algoritma klasifikasi seperti Naive Bayes dan K-Nearest Neighbor memungkinkan proses pengelompokan terhadap tingkat kepuasan pelanggan dilakukan secara lebih sistematis dan terukur. Dengan adanya model klasifikasi, pola

kecenderungan pelanggan dapat terlihat lebih jelas sehingga hasil analisis dapat digunakan untuk membandingkan performa kedua metode tersebut dalam menghasilkan prediksi yang paling akurat.

Data mining merupakan proses pengumpulan dan pengolahan data dengan tujuan mengekstrak informasi penting dalam data yang besar (Indah et al., 2023). Proses ini tidak hanya sekadar mengumpulkan data, tetapi juga melibatkan serangkaian tahapan terstruktur yang bertujuan untuk meningkatkan kualitas data sebelum dianalisis. Tahapan seperti pembersihan dan transformasi data berperan besar dalam memastikan bahwa data yang digunakan bebas dari kesalahan, konsisten, dan siap untuk diproses lebih lanjut. Dengan data yang berkualitas, algoritma data mining dapat bekerja secara optimal dalam menemukan pola, hubungan, serta kecenderungan yang tersembunyi di dalam data. Hal ini menjadikan data mining sebagai alat yang efektif dalam membantu organisasi atau institusi memahami kondisi yang sebenarnya berdasarkan data yang dimiliki.

Dalam penelitian ini, data mining adalah teknik yang digunakan untuk mengekstrak pengetahuan dari kumpulan data menggunakan teknik statistik dan matematika (S. A. Hasibuan et al., 2023). Penggunaan algoritma Naive Bayes dan K-Nearest Neighbor (KNN) memungkinkan proses pengelompokan data kepuasan pelanggan dilakukan secara sistematis dan terukur. Naive Bayes bekerja berdasarkan konsep probabilitas sehingga mampu memberikan hasil prediksi yang cepat dan efisien, sedangkan KNN mengklasifikasikan data berdasarkan kedekatan jarak antar data sehingga lebih mempertimbangkan kemiripan karakteristik pelanggan. Dengan membandingkan kedua metode tersebut, penelitian ini tidak hanya berfokus pada hasil klasifikasi, tetapi juga pada kinerja masing-masing algoritma dalam menghasilkan tingkat akurasi yang optimal. Hasil analisis yang diperoleh diharapkan mampu memberikan gambaran yang lebih

jelas mengenai pola kepuasan pelanggan serta menjadi dasar yang kuat dalam pengambilan keputusan untuk meningkatkan kualitas layanan secara berkelanjutan.

### **3. Model Klasifikasi**

Model klasifikasi memiliki beberapa kekurangan yang dapat memengaruhi kualitas hasil prediksi, terutama ketika diterapkan pada data yang kompleks dan memiliki banyak variabel. Salah satu kelemahannya adalah sensitivitas terhadap kualitas dan jumlah data, karena model tertentu dapat menghasilkan prediksi yang tidak stabil ketika data yang digunakan tidak seimbang, terdapat noise, atau jumlahnya terlalu sedikit. Selain itu, beberapa model tidak mampu menangkap hubungan non-linear antar variabel sehingga hasil klasifikasi menjadi kurang akurat. Keterbatasan lain muncul ketika model tidak mampu beradaptasi dengan data baru atau kondisi yang berubah, sehingga performanya menurun seiring waktu. Proses pemilihan parameter yang tidak tepat, keterbatasan dalam generalisasi, serta risiko overfitting atau underfitting juga menjadi bagian dari kekurangan umum dalam penggunaan model klasifikasi.

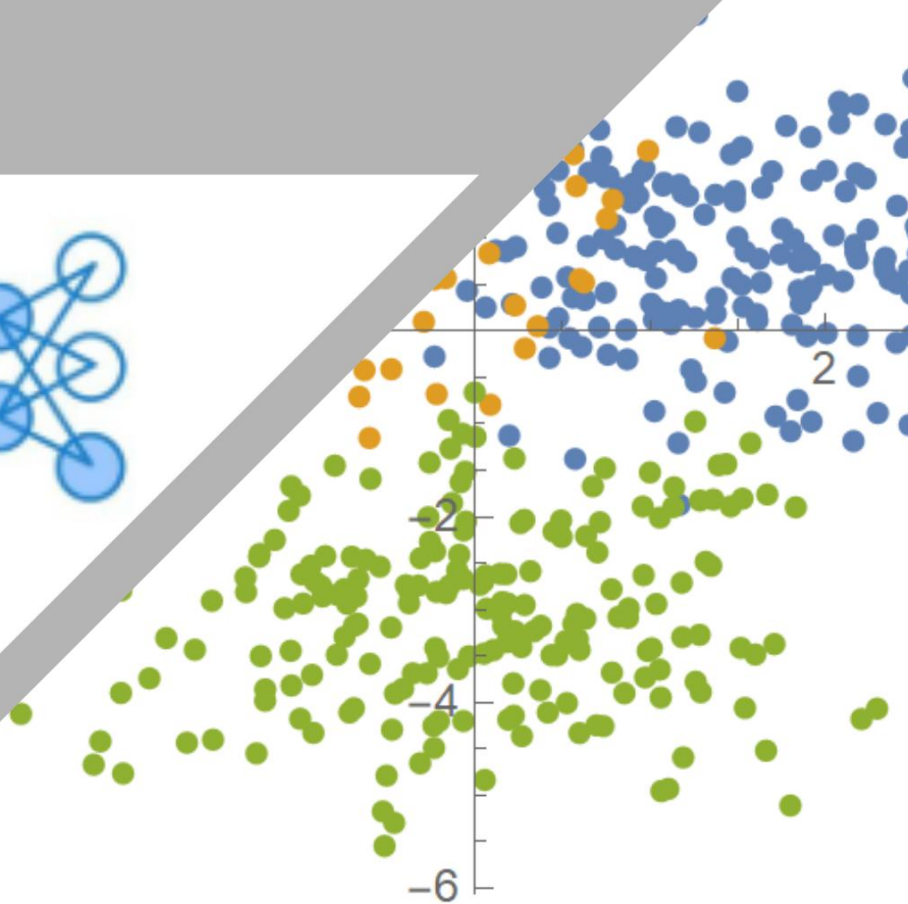
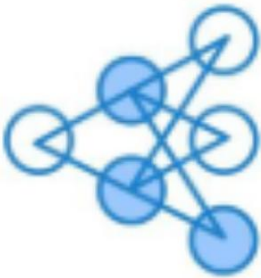
Kekurangan tersebut menjadi alasan pentingnya proses evaluasi yang komprehensif ketika menerapkan algoritma klasifikasi untuk mengolah data kepuasan pelanggan di Mie Gacoan Rantauprapat. Data yang beragam, melibatkan karakteristik pelayanan, kualitas makanan, hingga kondisi tempat, menuntut model yang mampu beradaptasi dengan pola yang tidak selalu konsisten. Variasi persepsi pelanggan dapat menyebabkan noise dan ketidakteraturan pada data, sehingga algoritma yang terlalu bergantung pada asumsi tertentu berisiko menghasilkan prediksi yang meleset.

Kekurangan yang dimiliki oleh model klasifikasi dapat memberikan dampak signifikan terhadap kualitas hasil prediksi, terutama ketika model diterapkan pada data yang kompleks dan memiliki banyak variabel. Sensitivitas terhadap kualitas dan jumlah data menjadi salah satu permasalahan utama, karena data yang tidak seimbang, mengandung noise, atau berjumlah terbatas dapat menyebabkan hasil klasifikasi menjadi kurang stabil. Selain itu, beberapa model klasifikasi memiliki keterbatasan dalam menangkap hubungan non-linear antar variabel, sehingga tidak mampu merepresentasikan pola data secara menyeluruh. Permasalahan lain muncul ketika model tidak mampu beradaptasi dengan perubahan data atau kondisi baru, yang mengakibatkan penurunan performa seiring waktu. Pemilihan parameter yang kurang tepat, keterbatasan kemampuan generalisasi, serta risiko terjadinya overfitting atau underfitting juga menjadi faktor yang dapat mengurangi keakuratan hasil klasifikasi.

Dalam konteks klasifikasi kepuasan pelanggan di Mie Gacoan Rantauprapat, kekurangan tersebut perlu menjadi perhatian penting dalam proses analisis dan evaluasi model. Data kepuasan pelanggan yang beragam, mencakup aspek pelayanan, kualitas makanan, harga, dan kenyamanan tempat, memiliki karakteristik yang tidak selalu konsisten dan cenderung subjektif. Variasi persepsi pelanggan berpotensi menimbulkan noise dan ketidakteraturan pada data, sehingga algoritma yang terlalu bergantung pada asumsi tertentu dapat menghasilkan prediksi yang kurang tepat. Oleh karena itu, diperlukan evaluasi kinerja yang komprehensif serta perbandingan antar metode klasifikasi untuk memastikan bahwa model yang digunakan mampu memberikan hasil yang optimal dan dapat diandalkan sebagai dasar pengambilan keputusan.



# BAB 2





## BAB 2

# ALGORITMA NAIVE BAYES

### 1. Pengertian Naive Bayes

Naive Bayes merupakan metode klasifikasi probabilitas sederhana yang berdasarkan pada teorema Bayes dengan asumsi bahwa semua atribut bersifat independen atau tidak berhubungan (Punkastyo et al., 2024). Meskipun asumsi independensi tersebut sering kali tidak sepenuhnya terpenuhi pada data dunia nyata, metode ini tetap terbukti efektif karena proses perhitungannya sederhana, cepat, serta mampu menangani jumlah data yang besar. Naive Bayes menghitung probabilitas posterior suatu kelas berdasarkan kombinasi probabilitas prior dan likelihood dari setiap fitur, lalu memilih kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi. Keunggulan lain dari metode ini adalah kemampuannya menangani data kategorikal maupun numerik, sensitivitas yang rendah terhadap fitur yang tidak relevan, serta performa yang stabil meskipun jumlah data pelatihan terbatas.

Rumus Teorema Bayes:

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

Keterangan:

C : Kelas (misalnya puas, tidak puas)

X : Data atau atribut yang diamati

$P(C | X)$  : Probabilitas data X termasuk ke kelas C

$P(C)$  : Probabilitas awal (prior) kelas C

$P(X | C)$  : Probabilitas data X muncul pada kelas C

$P(X)$  : Probabilitas data X

Naive Bayes merupakan metode klasifikasi sederhana untuk menghitung probabilitas kombinasi data tertentu (Siregar et al., 2023). Algoritma ini memungkinkan proses identifikasi pola penilaian pelanggan secara lebih sistematis dengan menghitung peluang kecenderungan setiap kategori kepuasan. Melalui pendekatan perhitungan probabilitas tersebut, model mampu memberikan gambaran kecenderungan tingkat kepuasan yang muncul dari data yang tersedia, sehingga hasil klasifikasi dapat digunakan sebagai dasar evaluasi metode dan perbandingan performa dengan algoritma lain seperti K-Nearest Neighbor yang memiliki karakteristik perhitungan berbeda.

Metode Naive Bayes adalah metode klasifikasi yang akan mengklasifikasikan data atau teks ke dalam kategori terdekat (F. F. Hasibuan et al., 2023). Naive Bayes bekerja dengan mengasumsikan bahwa setiap atribut dalam data bersifat independen satu sama lain, sehingga proses perhitungan menjadi lebih sederhana dan efisien. Teorema Bayes adalah model matematika yang didasarkan pada statistik dan probabilitas (Alam et al., 2023). Dalam konteks penelitian kepuasan pelanggan, Naive Bayes memungkinkan peneliti untuk mengidentifikasi kecenderungan tingkat kepuasan pelanggan berdasarkan pola penilaian yang muncul dari data historis, sehingga hasil klasifikasi dapat menggambarkan kondisi kepuasan secara objektif.

Metode Naive Bayes adalah metode yang digunakan untuk melakukan klasifikasi sederhana dengan menghitung probabilitas kombinasi dari data tertentu. Dengan menghitung probabilitas setiap kelas kepuasan (Nasution et al., 2023). Naive Bayes telah terbukti efektif dalam berbagai tugas klasifikasi teks (Apriyani et al., 2023), algoritma ini mampu memberikan prediksi yang jelas dan terukur terhadap data baru. Selain itu, hasil klasifikasi yang diperoleh dapat dibandingkan secara langsung dengan metode lain seperti K-Nearest Neighbor (KNN) untuk mengetahui

perbedaan tingkat akurasi dan efektivitas masing-masing metode. Melalui perbandingan tersebut, penelitian ini diharapkan dapat memberikan gambaran yang komprehensif mengenai keunggulan dan keterbatasan algoritma Naive Bayes dalam mengklasifikasikan tingkat kepuasan pelanggan, sekaligus menjadi referensi dalam pemilihan metode yang paling sesuai untuk analisis data kepuasan di masa mendatang.

## **2. Konsep Dasar Metode Naive Bayes**

Naive Bayes merupakan salah satu metode yang dapat digunakan untuk mengklasifikasikan data. Prinsip dasarnya adalah menghitung peluang suatu data masuk ke kelas tertentu berdasarkan frekuensi kemunculan fitur-fitur yang menyusunnya. Pendekatan probabilistik ini memungkinkan proses perhitungan berjalan sangat efisien meskipun jumlah data besar, karena model hanya membutuhkan informasi statistik sederhana seperti probabilitas prior dan probabilitas kondisi. Naive Bayes juga dikenal memiliki kinerja yang baik pada data kategorikal, stabil meskipun terdapat noise, serta mampu menghasilkan prediksi yang cepat tanpa membutuhkan proses pelatihan yang kompleks. Sifatnya yang ringan dan interpretatif menjadikan algoritma ini sebagai salah satu metode klasifikasi yang banyak digunakan di berbagai bidang, mulai dari analisis teks, deteksi spam, hingga prediksi berbasis data pelanggan.

Metode Naive Bayes didasarkan pada teorema probabilitas Bayes yang menggambarkan hubungan antara probabilitas kondisional dari beberapa peristiwa. Algoritma ini dapat mengidentifikasi pola umum dari respon pelanggan dengan memetakan frekuensi penilaian tertentu pada kategori kepuasan, sehingga model mampu membedakan mana atribut yang paling berpengaruh terhadap suatu kelas kepuasan. Hubungan antar variabel yang dikonversi menjadi peluang membuat Naive Bayes

dapat memberikan gambaran probabilistik mengenai tingkat kepuasan pelanggan baru berdasarkan data historis.

Metode Naïve Bayes memiliki peranan yang signifikan dalam proses klasifikasi data kepuasan pelanggan karena kemampuannya dalam mengolah data secara efisien dan terstruktur. Pendekatan probabilistik yang digunakan memungkinkan algoritma ini untuk mengidentifikasi pola umum dari respon pelanggan dengan menghitung peluang setiap atribut terhadap kelas kepuasan tertentu, sehingga hasil prediksi yang dihasilkan bersifat objektif dan terukur. Selain itu, asumsi independensi antar atribut menjadikan proses perhitungan lebih sederhana tanpa mengurangi efektivitas model secara signifikan, khususnya pada data kategorikal yang umum digunakan dalam penilaian kepuasan. Kemampuan Naïve Bayes dalam menangani data berukuran besar serta toleransinya terhadap noise menjadikannya sangat sesuai untuk diterapkan pada penelitian ini. Dengan memanfaatkan data historis pelanggan, model Naïve Bayes mampu memberikan gambaran probabilistik mengenai tingkat kepuasan pelanggan baru, sehingga hasil klasifikasi yang diperoleh dapat digunakan sebagai dasar evaluasi kualitas layanan dan menjadi bahan pertimbangan dalam pengambilan keputusan berbasis data.

### **3. Kelebihan Metode Naive Bayes**

Naive Bayes merupakan metode probabilistik pengklasifikasian sederhana berdasarkan Teorema Bayes dimana pengklasifikasian dilakukan melalui training set sejumlah data secara efisien (Harahap et al., 2023). Salah satu keunggulan utamanya adalah kemampuannya bekerja dengan baik meskipun jumlah data pelatihan relatif sedikit, karena metode ini hanya membutuhkan estimasi probabilitas sederhana berdasarkan frekuensi kemunculan data. Asumsi independensi antar fitur juga membuat proses komputasi menjadi jauh lebih ringan, sehingga

perhitungan dapat dilakukan dengan cepat meskipun data berukuran besar. Selain itu, Naive Bayes cukup tangguh terhadap data yang tidak lengkap serta mampu menangani fitur kategorikal maupun numerik. Pendekatan probabilistik yang digunakan juga memberikan interpretasi yang jelas mengenai peluang suatu data termasuk dalam kelas tertentu, sehingga mempermudah analisis dan pengambilan keputusan berbasis hasil klasifikasi.

Metode Naive Bayes Classifier merupakan salah satu pengklasifikasi statistik, dimana klasifikasi ini dapat memprediksi probabilitas keanggotaan kelas suatu data yang akan masuk ke dalam kelas tertentu, sesuai dengan perhitungan probabilitas. Metode ini mampu memproses data seperti penilaian layanan, kecepatan penyajian, rasa makanan, atau kenyamanan secara efisien tanpa memerlukan struktur data yang kompleks. Kecepatan komputasi Naive Bayes memungkinkan proses pelatihan dan pengujian dilakukan secara cepat meskipun dataset berisi banyak entri pelanggan, sehingga hasil klasifikasi dapat diperoleh lebih efektif. Selain itu, sifat model yang mudah diinterpretasikan membantu menggambarkan pengaruh setiap variabel terhadap pembentukan kategori kepuasan, sehingga hasil analisis lebih mudah dipahami dan dapat digunakan sebagai dasar peningkatan kualitas layanan.

Naive Bayes merupakan metode klasifikasi yang sangat relevan untuk digunakan dalam analisis kepuasan pelanggan karena kesederhanaan dan efisiensi yang dimilikinya. Algoritma ini mampu bekerja dengan baik meskipun jumlah data pelatihan relatif terbatas, karena hanya bergantung pada perhitungan probabilitas berdasarkan frekuensi kemunculan data. Asumsi independensi antar fitur menjadikan proses komputasi lebih ringan dan cepat, sehingga sangat sesuai diterapkan pada dataset berukuran besar. Selain itu, Naive Bayes memiliki kemampuan yang baik dalam menangani data yang tidak lengkap serta dapat

digunakan pada data kategorikal maupun numerik, seperti penilaian layanan, kecepatan penyajian, rasa makanan, dan kenyamanan. Model probabilistik yang dihasilkan juga memberikan interpretasi yang jelas mengenai peluang suatu data termasuk ke dalam kelas kepuasan tertentu. Dengan karakteristik tersebut, Naive Bayes tidak hanya efektif dalam menghasilkan prediksi yang akurat, tetapi juga memudahkan peneliti dalam memahami pola kepuasan pelanggan dan menjadikannya sebagai dasar pengambilan keputusan untuk peningkatan kualitas layanan.

#### **4. Kekurangan Metode Naive Bayes**

Naive Bayes merupakan metode pengklasifikasian yang paling terkenal dengan menggunakan probabilitas yang sederhana (Rozi, 2023). Algoritma ini mengharuskan setiap fitur dianggap saling independen, padahal dalam banyak kasus dunia nyata hubungan antarvariabel justru saling bergantung. Ketika asumsi independensi tersebut tidak terpenuhi, akurasi model dapat menurun karena perhitungan probabilitas menjadi tidak sepenuhnya mencerminkan kondisi asli data. Selain itu, Naive Bayes kurang optimal ketika menghadapi data dengan distribusi yang tidak seimbang, terutama jika terdapat kelas dengan jumlah sampel yang jauh lebih sedikit dibandingkan kelas lain. Model juga cenderung kesulitan menangani fitur numerik yang tidak mengikuti distribusi normal, serta sensitif terhadap data yang mengandung noise atau korelasi kuat antarfitur.

Naive Bayes adalah Teorema Bayes, yaitu melakukan klasifikasi dengan melakukan perhitungan nilai probabilitas. Keterbatasan tersebut berdampak pada proses klasifikasi kepuasan pelanggan karena data pelanggan biasanya memiliki hubungan antarvariabel yang kompleks, seperti hubungan antara kualitas layanan, waktu penyajian, dan harga. Variabel-variabel ini sering menunjukkan korelasi, sehingga asumsi independensi Naive

Bayes dapat menyebabkan model salah mengestimasi probabilitas suatu kelas. Selain itu, jika terdapat kategori penilaian pelanggan yang jumlahnya jauh lebih kecil daripada kategori lainnya, model dapat cenderung mengabaikan kelas minoritas meskipun informasi tersebut penting dalam menggambarkan variasi kepuasan.

Naive Bayes merupakan metode klasifikasi berbasis probabilitas yang memiliki sejumlah keterbatasan dalam penerapannya pada data dunia nyata. Asumsi independensi antar fitur yang digunakan oleh algoritma ini sering kali tidak sesuai dengan kondisi data sebenarnya, karena variabel-variabel dalam data kepuasan pelanggan umumnya saling berkaitan. Ketika hubungan antarvariabel tersebut diabaikan, hasil perhitungan probabilitas dapat menjadi kurang akurat dan tidak sepenuhnya merepresentasikan kondisi data yang sebenarnya. Selain itu, Naive Bayes juga kurang optimal dalam menangani data dengan distribusi kelas yang tidak seimbang, di mana kelas dengan jumlah data lebih sedikit cenderung terabaikan oleh model.

Keterbatasan lainnya terlihat pada kemampuan Naive Bayes dalam menangani fitur numerik yang tidak mengikuti distribusi tertentu serta sensitivitasnya terhadap data yang mengandung noise atau korelasi yang kuat. Dalam konteks klasifikasi kepuasan pelanggan, kondisi ini dapat memengaruhi ketepatan model dalam membedakan tingkat kepuasan yang memiliki karakteristik serupa. Hubungan kompleks antara kualitas layanan, kecepatan penyajian, harga, dan faktor lainnya menuntut model yang mampu menangkap keterkaitan antarvariabel secara lebih baik. Oleh karena itu, meskipun Naive Bayes memiliki keunggulan dari sisi kesederhanaan dan kecepatan, keterbatasan tersebut perlu diperhatikan agar hasil klasifikasi yang dihasilkan dapat dievaluasi secara objektif dan dibandingkan dengan metode lain.

## **5. Pengolahan Data Metode Naive Bayes**

Metode Naive Bayes merupakan salah satu algoritma klasifikasi dalam data mining yang bekerja berdasarkan konsep probabilitas Bayes. Metode ini digunakan untuk menentukan kemungkinan suatu data termasuk ke dalam kelas tertentu berdasarkan atribut-atribut yang dimilikinya. Dalam penelitian kepuasan pelanggan, Naive Bayes dapat dimanfaatkan untuk mengelompokkan pelanggan ke dalam kategori puas dan tidak puas berdasarkan beberapa variabel penilaian, seperti kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Keunggulan metode ini terletak pada proses perhitungannya yang sederhana, cepat, dan efektif untuk data kategorikal.

Pengolahan data dengan metode Naive Bayes diawali dari proses pengumpulan data. Data penelitian umumnya diperoleh melalui penyebaran kuesioner kepada responden, baik secara langsung maupun menggunakan media daring seperti Google Form. Dalam konteks penelitian kepuasan pelanggan, data yang dikumpulkan berisi penilaian responden terhadap atribut layanan yang telah ditentukan. Data tersebut kemudian diinput dan direkap secara sistematis agar dapat digunakan sebagai bahan analisis. Tahap ini sangat penting karena kualitas data awal sangat menentukan kualitas model klasifikasi yang akan dihasilkan.

Setelah data terkumpul, dilakukan tahap pembersihan data (data cleaning) untuk memastikan bahwa dataset berada dalam kondisi yang layak dianalisis. Pada tahap ini, peneliti memeriksa kemungkinan adanya data ganda, data tidak lengkap, kesalahan penulisan, atau ketidaksesuaian isi jawaban. Data yang tidak valid dapat diperbaiki atau dikeluarkan dari dataset agar tidak menimbulkan bias dalam proses klasifikasi. Dengan demikian, data yang digunakan benar-benar mewakili kondisi responden dan dapat mendukung perhitungan probabilitas secara akurat.

Tahap berikutnya adalah pembagian data menjadi data training dan data testing. Data training digunakan sebagai dasar pembentukan model Naive Bayes, sedangkan data testing digunakan untuk menguji kemampuan model dalam melakukan klasifikasi pada data baru. Melalui data training, sistem menghitung pola kemunculan atribut pada masing-masing kelas kepuasan. Sementara itu, data testing berfungsi untuk melihat sejauh mana model mampu memprediksi kategori puas atau tidak puas secara objektif. Pembagian data ini penting agar evaluasi model tidak hanya bergantung pada data yang sudah dipelajari sebelumnya.

Dalam metode Naive Bayes, proses inti dimulai dengan menghitung probabilitas awal (prior probability) dari setiap kelas. Probabilitas prior diperoleh dengan membandingkan jumlah data dalam masing-masing kelas dengan jumlah seluruh data training. Misalnya, jika dari 20 data training terdapat 13 data dengan kategori puas dan 7 data dengan kategori tidak puas, maka probabilitas awal kelas puas adalah  $13/20$  dan probabilitas awal kelas tidak puas adalah  $7/20$ . Probabilitas awal ini menjadi dasar untuk mengetahui kecenderungan umum distribusi kelas dalam dataset.

Selanjutnya dilakukan perhitungan probabilitas bersyarat (conditional probability) untuk setiap atribut pada masing-masing kelas. Perhitungan ini bertujuan untuk mengetahui seberapa besar kemungkinan suatu nilai atribut muncul pada kelas tertentu. Sebagai contoh, atribut baik pada kualitas pelayanan dapat memiliki peluang lebih besar pada kelas puas, sedangkan atribut kurang baik dapat memiliki peluang lebih besar pada kelas tidak puas. Proses perhitungan ini dilakukan untuk seluruh atribut yang digunakan dalam penelitian, sehingga diperoleh gambaran pola hubungan antara karakteristik layanan dengan tingkat kepuasan pelanggan.

Metode Naive Bayes bekerja dengan asumsi bahwa setiap atribut bersifat independen satu sama lain. Artinya, kemunculan satu atribut dianggap tidak memengaruhi atribut lainnya dalam menentukan kelas. Walaupun asumsi ini tergolong sederhana, metode Naive Bayes tetap banyak digunakan karena mampu memberikan hasil klasifikasi yang baik pada berbagai kasus. Dalam proses klasifikasi, probabilitas setiap atribut terhadap masing-masing kelas dikalikan dengan probabilitas prior, sehingga diperoleh probabilitas akhir untuk setiap kelas. Kelas dengan nilai probabilitas terbesar kemudian ditetapkan sebagai hasil klasifikasi data tersebut.

Tahap akhir dari pengolahan data menggunakan Naive Bayes adalah penentuan hasil klasifikasi dan evaluasi model. Setiap data uji akan dikategorikan ke dalam kelas puas atau tidak puas berdasarkan probabilitas akhir yang tertinggi. Hasil klasifikasi tersebut selanjutnya dapat dianalisis untuk melihat pola kepuasan pelanggan secara umum dan untuk mengukur kinerja model. Dengan demikian, pengolahan data menggunakan metode Naive Bayes tidak hanya membantu mengelompokkan data secara sistematis, tetapi juga memberikan dasar analitis yang kuat dalam memahami faktor-faktor yang memengaruhi kepuasan pelanggan. Dalam penelitian bidang kuliner, metode ini dapat menjadi alat yang efektif untuk mendukung evaluasi layanan dan pengambilan keputusan berbasis data.

# BAB 3





## BAB 3

### ALGORITMA K-NEAREST NEIGHBOR (KNN)

#### 1. Pengertian K-Nearest Neighbor

Algoritma klasifikasi K-Nearest Neighbor adalah model yang dapat digunakan untuk mengklasifikasikan data berdasarkan data yang paling dekat dengan objek yang diinginkan. KNN termasuk ke dalam algoritma lazy learner, artinya metode ini tidak membentuk model secara eksplisit, melainkan menyimpan seluruh data latih dan melakukan proses klasifikasi ketika ada data baru yang ingin diprediksi. Proses penentuan kelas dilakukan dengan menghitung jarak biasanya menggunakan Euclidean distance antara data baru dengan sejumlah tetangga terdekat yang ditentukan oleh nilai K. Data baru kemudian diklasifikasikan berdasarkan mayoritas kelas dari tetangga-tetangga terdekat tersebut. KNN adalah contoh algoritma berbasis pembelajaran, di mana kumpulan data pelatihan (training) disimpan, sehingga klasifikasi untuk data baru yang belum terklasifikasi diperoleh dengan membandingkan data yang paling mirip dengan data pelatihan (Adjani et al., 2023).

Rumus Jarak Euclidean:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

- $d(x,y)$  = jarak antara data uji dan data latih
- $x_i$  = nilai atribut ke- dari data uji
- $y_i$  = nilai atribut ke- dari data latih
- $n$  = jumlah atribut

Metode K-Nearest Neighbor ini merupakan teknik yang digunakan dalam penambangan data dan pembelajaran mesin untuk melakukan klasifikasi dan regresi berdasarkan pengamatan tetangga terdekat dari data (Irmayani et al., 2023). Proses pengklasifikasian tingkat kepuasan dapat dilakukan dengan membandingkan karakteristik data pelanggan yang baru dengan riwayat penilaian pelanggan sebelumnya sehingga algoritma ini dapat mengidentifikasi kelompok kepuasan yang paling mendekati pola data tersebut. Ketika jumlah pelanggan besar dan pola penilaiannya bervariasi, KNN mampu mengakomodasi data secara fleksibel tanpa memerlukan asumsi distribusi tertentu.

Metode K-Nearest Neighbor (KNN) merupakan salah satu algoritma machine learning yang dikenal karena kesederhanaan konsep dan kemudahan dalam implementasinya. Algoritma ini bekerja dengan cara mengklasifikasikan data baru berdasarkan kedekatan jaraknya dengan sejumlah data terdekat (neighbor) yang telah memiliki label kelas sebelumnya. Dalam konteks klasifikasi kepuasan pelanggan, KNN membandingkan karakteristik penilaian pelanggan baru, seperti kualitas layanan, kecepatan penyajian, dan kenyamanan, dengan data historis pelanggan yang sudah ada. Pendekatan ini memungkinkan KNN untuk mengidentifikasi kelompok kepuasan yang paling sesuai tanpa memerlukan proses pelatihan model yang kompleks. Selain itu, KNN tidak bergantung pada asumsi distribusi data tertentu, sehingga mampu menangani pola penilaian pelanggan yang beragam dan dinamis. Fleksibilitas tersebut menjadikan KNN efektif digunakan pada dataset dengan karakteristik yang bervariasi, terutama ketika jumlah data pelanggan cukup besar dan hubungan antarvariabel sulit dimodelkan secara statistik.

## **2. Konsep Dasar Metode KNN**

Metode K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi yang bekerja dengan prinsip kedekatan atau similaritas antar data. KNN menentukan kelas suatu data baru dengan cara mencari  $k$  tetangga terdekat berdasarkan jarak, biasanya menggunakan Euclidean Distance. Algoritma ini termasuk metode lazy learning karena tidak memerlukan proses pelatihan model secara eksplisit; seluruh data latih disimpan dan digunakan saat proses klasifikasi berlangsung. Kinerja KNN sangat dipengaruhi oleh pemilihan nilai  $k$ , kualitas data, serta distribusi fitur pada dataset. Semakin rendah nilai  $k$ , hasil klasifikasi menjadi lebih sensitif terhadap data yang tidak normal, sedangkan nilai  $k$  yang terlalu besar dapat mengaburkan batas antar kelas. Keunggulan KNN terletak pada kesederhanaannya dan kemampuannya beradaptasi dengan pola data non-linier, namun algoritma ini cukup rentan terhadap data bising, perbedaan skala antar fitur, dan jumlah data yang besar karena proses perhitungan jarak dilakukan berulang.

Penerapan KNN dalam penelitian kepuasan pelanggan memberi peluang untuk memperoleh hasil klasifikasi yang lebih intuitif karena metode ini mengelompokkan pelanggan berdasarkan kemiripan karakteristik penilaian. Ketika data pelanggan menunjukkan pola yang beragam baik dari aspek pelayanan, kualitas makanan, harga, maupun kenyamanan KNN dapat mengidentifikasi kecenderungan kelompok tertentu melalui perhitungan jarak dengan data sebelumnya. Proses pencarian tetangga terdekat membantu memperlihatkan bagaimana pelanggan dengan karakteristik serupa cenderung memberikan tingkat kepuasan yang sama. Hasil klasifikasi tersebut dapat digunakan untuk melihat segmentasi penilaian pelanggan, membandingkan performa metode KNN dengan metode lainnya.

Metode K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi berbasis kedekatan yang menentukan kelas suatu data baru berdasarkan kemiripan dengan data lain yang telah memiliki label kelas. Algoritma ini bekerja dengan menghitung jarak antar data, umumnya menggunakan Euclidean Distance, untuk menemukan sejumlah tetangga terdekat yang paling relevan. KNN termasuk ke dalam metode lazy learning karena tidak membangun model pelatihan secara khusus, melainkan menyimpan seluruh data latih dan menggunakannya langsung saat proses klasifikasi berlangsung. Kinerja algoritma ini sangat dipengaruhi oleh pemilihan nilai  $k$ , kualitas data, serta distribusi fitur yang digunakan, di mana nilai  $k$  yang terlalu kecil dapat menyebabkan hasil klasifikasi menjadi sensitif terhadap noise, sedangkan nilai  $k$  yang terlalu besar berpotensi mengaburkan batas antar kelas. Keunggulan KNN terletak pada kesederhanaan dan kemampuannya menangani pola data non-linier, namun algoritma ini memiliki kelemahan berupa ketergantungan pada skala data, sensitivitas terhadap data bising, serta meningkatnya beban komputasi ketika jumlah data semakin besar.

Penerapan KNN dalam penelitian kepuasan pelanggan memungkinkan proses klasifikasi dilakukan secara lebih intuitif karena pengelompokan didasarkan pada kemiripan karakteristik penilaian pelanggan. Algoritma ini mampu mengidentifikasi kecenderungan tingkat kepuasan pelanggan dengan membandingkan data penilaian baru terhadap data historis yang memiliki karakteristik serupa. Ketika pelanggan menunjukkan variasi penilaian pada aspek pelayanan, kualitas produk, harga, dan kenyamanan, KNN dapat memetakan pola kepuasan berdasarkan kedekatan nilai atribut tersebut. Proses pencarian tetangga terdekat membantu menunjukkan bahwa pelanggan dengan karakteristik yang mirip cenderung memberikan tingkat kepuasan yang sama. Hasil klasifikasi yang diperoleh dapat

dimanfaatkan untuk melihat segmentasi kepuasan pelanggan serta menjadi dasar dalam membandingkan kinerja metode KNN dengan algoritma klasifikasi lainnya.

### **3. Kelebihan Metode KNN**

Metode K-Nearest Neighbor (KNN) memiliki beberapa kelebihan yang membuatnya banyak digunakan dalam permasalahan klasifikasi dan pengenalan pola. KNN bersifat sederhana namun efektif karena tidak memerlukan proses pelatihan model secara eksplisit, algoritma ini bekerja dengan membandingkan data baru terhadap data yang sudah ada berdasarkan kedekatan jarak. Sifat non-parametriknya membuat KNN mampu beradaptasi dengan berbagai bentuk distribusi data tanpa asumsi khusus, sehingga performanya tetap stabil meski data memiliki pola yang kompleks. Selain itu, KNN sangat intuitif dan mudah diimplementasikan, sementara tingkat akurasi dapat ditingkatkan dengan penentuan nilai  $k$  yang tepat serta pemilihan metode pengukuran jarak yang sesuai. Fleksibilitas ini menjadikan KNN unggul saat digunakan pada dataset yang bersifat dinamis dan memiliki variabel-variabel yang beragam.

Penerapan KNN pada analisis kepuasan pelanggan memberikan kelebihan karena algoritma ini mampu mengelompokkan data berdasarkan kemiripan perilaku atau karakteristik pelanggan sebelumnya. Kedekatan jarak antar data membantu memprediksi kategori kepuasan secara lebih akurat terutama ketika pola penilaian pelanggan cenderung menunjukkan kesamaan di antara kelompok tertentu. Algoritma ini juga mempermudah proses interpretasi hasil karena keputusan klasifikasi dapat ditelusuri langsung dari data tetangga terdekat, sehingga memudahkan evaluasi keandalan prediksi yang dihasilkan. Keunggulan tersebut membuat KNN relevan digunakan untuk membandingkan efektivitas metode klasifikasi lainnya.

Metode K-Nearest Neighbor (KNN) memiliki berbagai keunggulan yang menjadikannya banyak digunakan dalam permasalahan klasifikasi dan pengenalan pola. Algoritma ini bersifat sederhana namun tetap efektif karena tidak memerlukan proses pelatihan model secara eksplisit, melainkan bekerja dengan membandingkan data baru terhadap data yang sudah ada berdasarkan kedekatan jarak. Sifat non-parametrik yang dimiliki KNN membuatnya mampu beradaptasi dengan berbagai bentuk distribusi data tanpa memerlukan asumsi tertentu, sehingga tetap dapat bekerja dengan baik pada data yang memiliki pola kompleks. Selain itu, KNN mudah diimplementasikan dan dipahami, serta tingkat akurasi dapat ditingkatkan melalui pemilihan nilai  $k$  dan metode pengukuran jarak yang tepat. Fleksibilitas tersebut menjadikan KNN sangat sesuai digunakan pada dataset yang bersifat dinamis dan memiliki karakteristik data yang beragam.

Dalam analisis kepuasan pelanggan, penerapan KNN memberikan keuntungan karena algoritma ini mampu mengelompokkan data berdasarkan kemiripan perilaku dan karakteristik pelanggan. Kedekatan jarak antar data memungkinkan model untuk memprediksi kategori kepuasan dengan lebih akurat, terutama ketika pola penilaian pelanggan menunjukkan kesamaan dalam kelompok tertentu. Selain itu, proses klasifikasi KNN relatif mudah diinterpretasikan karena keputusan yang dihasilkan dapat ditelusuri langsung dari data tetangga terdekat yang memengaruhinya. Hal ini memudahkan peneliti dalam mengevaluasi keandalan hasil klasifikasi serta memahami pola kepuasan pelanggan secara lebih jelas. Keunggulan-keunggulan tersebut menjadikan KNN relevan untuk digunakan dan dibandingkan dengan metode klasifikasi lain dalam penelitian ini.

#### **4. Kekurangan Metode KNN**

Metode K-Nearest Neighbor (KNN) memiliki beberapa kekurangan yang dapat memengaruhi kinerjanya dalam proses klasifikasi. KNN sangat bergantung pada perhitungan jarak sehingga performanya menurun ketika jumlah data besar atau memiliki dimensi yang tinggi. Proses pencarian tetangga terdekat memerlukan komputasi yang cukup berat karena setiap data uji harus dibandingkan dengan seluruh data latih. Selain itu, KNN sensitif terhadap fitur yang memiliki skala berbeda, sehingga tanpa normalisasi hasil klasifikasi dapat menjadi bias. Metode ini juga rentan terhadap noise karena satu data yang tidak representatif dapat memengaruhi hasil, terutama jika jumlah tetangga ( $k$ ) yang digunakan kecil. Ketergantungan pada distribusi data membuat KNN kurang stabil jika data tidak terkelompok dengan jelas atau terdapat outlier dalam jumlah signifikan.

Kekurangan KNN tersebut berpotensi memengaruhi akurasi ketika digunakan untuk mengklasifikasikan kepuasan pelanggan yang memiliki banyak variabel penilaian berbeda, seperti rasa makanan, harga, pelayanan, dan kenyamanan tempat. Data pelanggan biasanya memiliki keragaman preferensi dan variasi skor yang tinggi, sehingga keberadaan outlier atau nilai ekstrem dapat mengganggu hasil kedekatan antar data. Ketergantungan metode ini pada jumlah data juga membuat proses klasifikasi menjadi lebih lambat ketika volume penilaian pelanggan semakin besar, sehingga efisiensi komputasi bisa menurun. Variasi skala antar variabel seperti waktu tunggu yang berbentuk angka dan penilaian rasa yang berupa kategori numerik dapat menyebabkan jarak antar data menjadi tidak proporsional jika tidak dilakukan preprocessing yang tepat.

## **5. Pengolahan Data Metode KNN**

Metode K-Nearest Neighbor (KNN) merupakan salah satu algoritma klasifikasi dalam bidang data mining dan machine learning yang bekerja berdasarkan prinsip kedekatan jarak antar data. Algoritma ini mengklasifikasikan suatu data baru dengan melihat kemiripan terhadap sejumlah data yang telah diketahui kelasnya. Dalam proses klasifikasi, KNN tidak membangun model terlebih dahulu, tetapi langsung melakukan perbandingan antara data uji dengan data latih. Oleh karena itu, metode ini sering disebut sebagai *lazy learning algorithm*, karena proses pembelajaran terjadi pada saat proses klasifikasi dilakukan.

Dalam penerapan metode KNN, tahap awal yang penting adalah pengolahan dan persiapan data. Data yang digunakan dalam penelitian sering kali berbentuk kategorikal, seperti penilaian terhadap kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Namun, karena metode KNN menggunakan perhitungan jarak matematis, maka data tersebut perlu melalui proses transformasi data. Transformasi ini dilakukan dengan mengubah data kategorikal menjadi nilai numerik tertentu sehingga dapat diproses oleh algoritma KNN dalam menghitung tingkat kedekatan antar data.

Setelah proses transformasi data selesai, langkah selanjutnya adalah pembagian dataset menjadi dua bagian utama, yaitu data training dan data testing. Data training berfungsi sebagai data pembelajaran yang telah memiliki label kelas, sedangkan data testing digunakan untuk menguji kemampuan model dalam melakukan klasifikasi terhadap data baru. Pembagian data ini bertujuan untuk memastikan bahwa model klasifikasi mampu melakukan generalisasi terhadap data yang belum pernah diproses sebelumnya, sehingga hasil analisis yang diperoleh dapat lebih objektif dan akurat.

Tahap utama dalam metode KNN adalah perhitungan jarak antar data, yang umumnya menggunakan rumus Euclidean Distance. Perhitungan ini dilakukan dengan mencari selisih nilai antara setiap atribut pada data uji dan data training, kemudian selisih tersebut dikuadratkan, dijumlahkan, dan diambil nilai akarnya untuk memperoleh nilai jarak. Nilai jarak ini menunjukkan tingkat kemiripan antara dua data, di mana semakin kecil nilai jarak yang dihasilkan maka semakin mirip karakteristik kedua data tersebut. Setelah jarak dihitung, sejumlah data terdekat ( $k$  tetangga terdekat) dipilih sebagai dasar untuk menentukan kelas dari data uji.

Berdasarkan proses tersebut, metode KNN mampu mengelompokkan data ke dalam kategori tertentu dengan mempertimbangkan kedekatan karakteristik antar data. Dalam penelitian mengenai kepuasan pelanggan, metode ini dapat digunakan untuk mengklasifikasikan pelanggan ke dalam kategori puas atau tidak puas berdasarkan atribut seperti kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Dengan memanfaatkan perhitungan jarak antar data, KNN dapat membantu mengidentifikasi pola kepuasan pelanggan secara lebih sistematis sehingga dapat digunakan sebagai dasar dalam evaluasi dan peningkatan kualitas layanan.



# BAB 4



**RapidM**

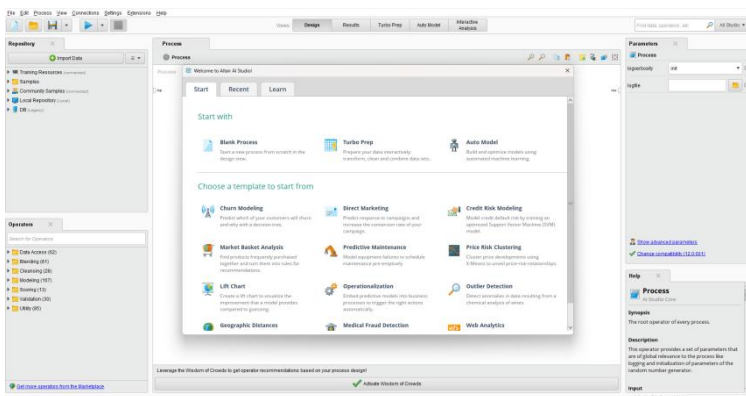


## BAB 4

# TOOLS PENGUJIAN ALGORITMA NAIVE BAYES DAN KNN

### 1. RapidMiner

RapidMiner merupakan salah satu perangkat lunak Data mining yang menyediakan berbagai fitur lengkap untuk proses analisis data, termasuk pembersihan data, transformasi, pemodelan, evaluasi, dan visualisasi. Aplikasi ini memiliki antarmuka berbasis drag-and-drop sehingga memudahkan pengguna dalam membangun alur proses tanpa harus menulis kode secara manual. RapidMiner juga mendukung banyak algoritma seperti klasifikasi, clustering, regresi, asosiasi, dan optimasi model, serta menyediakan berbagai operator yang dapat dikombinasikan sesuai kebutuhan analisis. Kemampuannya dalam menangani dataset besar, integrasi dengan berbagai format data, serta fitur evaluasi model yang detail menjadikannya alat yang banyak digunakan dalam riset akademik maupun industri.



**Gambar 2.** Aplikasi RapidMiner Studio (Altair Ai Studio 2026.0.1)

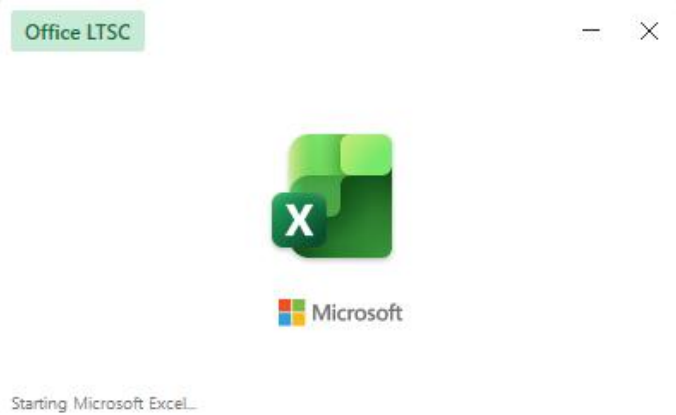
Penggunaan RapidMiner membantu mempercepat proses pemodelan karena alur kerja dapat dibangun secara sistematis,

mulai dari import data, preprocessing, penerapan algoritma Naive Bayes dan KNN, hingga evaluasi model menggunakan metrik akurasi. Seluruh tahapan analisis dapat divisualisasikan sehingga mempermudah pemahaman terhadap hasil yang diperoleh dan memastikan setiap langkah berjalan sesuai prosedur. Fitur perbandingan model memungkinkan pengguna menilai performa kedua algoritma secara objektif berdasarkan output yang dihasilkan.

Penggunaan RapidMiner memberikan kemudahan dan efisiensi dalam proses pemodelan data mining karena seluruh alur kerja dapat disusun secara sistematis dan terstruktur. Proses analisis dimulai dari tahap impor data, dilanjutkan dengan preprocessing seperti pembersihan dan transformasi data, kemudian penerapan algoritma Naive Bayes dan K-Nearest Neighbor (KNN), hingga tahap evaluasi model menggunakan metrik kinerja seperti akurasi. Setiap tahapan dapat divisualisasikan dengan jelas sehingga mempermudah pengguna dalam memahami alur proses dan hasil yang diperoleh. Selain itu, fitur perbandingan model yang tersedia pada RapidMiner memungkinkan peneliti untuk menilai kinerja kedua algoritma secara objektif berdasarkan hasil klasifikasi yang dihasilkan.

## **2. Microsoft Excel**

Microsoft Excel merupakan aplikasi pengolah angka yang dikembangkan oleh Microsoft dan banyak digunakan untuk mengelola, menganalisis, serta menyajikan data dalam bentuk tabel dan grafik. Aplikasi ini menyediakan berbagai fitur seperti perhitungan otomatis menggunakan rumus dan fungsi, pengelolaan data dalam bentuk spreadsheet, serta kemampuan untuk menyortir dan memfilter data. Dengan antarmuka yang mudah digunakan, Microsoft Excel menjadi alat yang populer di berbagai bidang, termasuk pendidikan, bisnis, dan penelitian.



**Gambar 3.** Aplikasi Microsoft Excel

Dalam konteks penelitian, Microsoft Excel sering dimanfaatkan sebagai alat bantu untuk pengolahan data awal sebelum dilakukan analisis lebih lanjut. Excel dapat digunakan untuk melakukan pembersihan data, transformasi data, serta perhitungan statistik sederhana. Selain itu, Excel juga mendukung pembuatan grafik dan visualisasi data yang membantu peneliti dalam memahami pola dan menyajikan hasil analisis secara lebih jelas dan sistematis.



# BAB 5





## **BAB 5**

# **IMPLEMENTASI ALGORITMA NAIVE BAYES DAN KNN PADA KLASIFIKASI KEPUASAN PELANGGAN**

### **1. Latar Belakang**

Pelayanan yang berkualitas menjadi aspek penting bagi setiap usaha kuliner dalam mempertahankan kepuasan serta loyalitas pelanggan. Mie Gacoan Rantauprapat sebagai salah satu restoran cepat saji yang banyak diminati, perlu memahami faktor-faktor yang memengaruhi kepuasan pelanggan agar dapat meningkatkan kualitas layanan dan produk yang diberikan. Dalam konteks ini, analisis terhadap data pelanggan menjadi langkah penting untuk mengetahui pola, kecenderungan, serta persepsi pelanggan terhadap pelayanan yang diterima. Pemanfaatan teknik Data Mining dapat membantu mengolah data tersebut sehingga menghasilkan informasi yang berguna bagi pengambilan keputusan.

Seiring meningkatnya jumlah pelanggan, data terkait penilaian kepuasan pun semakin banyak dan kompleks. Kondisi ini menuntut adanya metode analisis yang mampu mengolah data secara efektif dan menghasilkan keluaran yang akurat. Perkenalan terhadap teknik klasifikasi dalam Data Mining semakin relevan karena dapat memprediksi tingkat kepuasan berdasarkan variabel-variabel tertentu, seperti kualitas layanan, harga, kecepatan penyajian, dan kenyamanan tempat. Dengan demikian, penggunaan metode Data Mining dalam konteks ini bukan sekadar pilihan, melainkan kebutuhan untuk memahami perilaku pelanggan secara lebih mendalam.

Namun, perbedaan persepsi pelanggan serta beragamnya faktor yang memengaruhi kepuasan sering kali membuat proses klasifikasi menjadi kurang optimal jika hanya dilakukan secara manual. Munculnya kendala dalam menentukan model klasifikasi yang paling tepat menjadi tantangan tersendiri. Selain itu, setiap metode klasifikasi memiliki karakteristik dan tingkat akurasi yang berbeda sehingga perlu dilakukan analisis lebih lanjut untuk mengetahui metode mana yang paling sesuai dalam mengelompokkan kepuasan pelanggan di Mie Gacoan Rantauprapat. Ketidaktepatan dalam pemilihan metode berpotensi menghasilkan prediksi yang kurang akurat dan berdampak pada keputusan manajemen.

Kondisi tersebut menunjukkan perlunya penelitian yang membandingkan dua metode klasifikasi yang populer dalam Data Mining, yaitu Naive Bayes dan K-Nearest Neighbor (KNN). Keduanya sering digunakan dalam berbagai kasus klasifikasi, namun performanya dapat berbeda tergantung pada jenis data dan pola yang dianalisis. Penelitian sebelumnya menunjukkan hasil yang bervariasi antara kedua metode ini, sehingga masih diperlukan analisis lebih lanjut untuk mengetahui metode mana yang memberikan hasil lebih akurat dalam konteks kepuasan pelanggan pada sektor kuliner, khususnya di Mie Gacoan Rantauprapat.

Sebagai solusi dari permasalahan tersebut, penelitian ini menggunakan pendekatan Data Mining untuk mengklasifikasikan kepuasan pelanggan dengan membandingkan metode Naive Bayes dan KNN. Data Mining sendiri merupakan proses menggali pola atau informasi penting dari kumpulan data yang besar. Metode Naive Bayes bekerja berdasarkan prinsip probabilitas dan asumsi bahwa setiap fitur bersifat independen satu sama lain, sementara KNN menentukan klasifikasi berdasarkan kedekatan data baru dengan data sebelumnya. Dengan membandingkan

kedua metode ini, diharapkan penelitian mampu memberikan rekomendasi metode klasifikasi terbaik yang dapat membantu Mie Gacoan Rantauprapat meningkatkan kualitas layanan berdasarkan hasil analisis kepuasan pelanggan.

Berdasarkan latar belakang yang telah diuraikan sebelumnya, penelitian ini memfokuskan pada beberapa permasalahan utama yang berkaitan dengan analisis dan klasifikasi kepuasan pelanggan menggunakan pendekatan data mining. Permasalahan pertama berkaitan dengan bagaimana karakteristik dan pola data kepuasan pelanggan di Mie Gacoan Rantauprapat dapat diidentifikasi dan dianalisis sehingga data tersebut dapat dimanfaatkan sebagai dasar dalam proses klasifikasi menggunakan teknik data mining. Pemahaman terhadap karakteristik data menjadi tahap awal yang penting untuk mengetahui pola, atribut, serta hubungan antar variabel yang terdapat dalam data kepuasan pelanggan.

Permasalahan kedua berkaitan dengan proses penerapan metode klasifikasi, khususnya algoritma Naive Bayes dan K-Nearest Neighbor (KNN), dalam mengolah serta mengklasifikasikan data kepuasan pelanggan yang tersedia. Pada tahap ini, penelitian berfokus pada bagaimana kedua metode tersebut diterapkan dalam proses pengolahan data, mulai dari tahap persiapan data, proses pelatihan model, hingga proses klasifikasi terhadap data kepuasan pelanggan.

Permasalahan ketiga berkaitan dengan perbandingan kinerja kedua metode klasifikasi, yaitu Naive Bayes dan KNN, dalam menentukan tingkat kepuasan pelanggan di Mie Gacoan Rantauprapat. Perbandingan ini dilakukan dengan melihat tingkat akurasi yang dihasilkan oleh masing-masing metode dalam proses klasifikasi. Melalui perbandingan tersebut, diharapkan dapat diketahui metode mana yang memberikan hasil klasifikasi yang lebih optimal sehingga dapat digunakan sebagai pendekatan yang efektif dalam menganalisis kepuasan pelanggan.

## **2. Gambaran Umum Objek Penelitian**

### **a. Profil Perusahaan / Instansi**

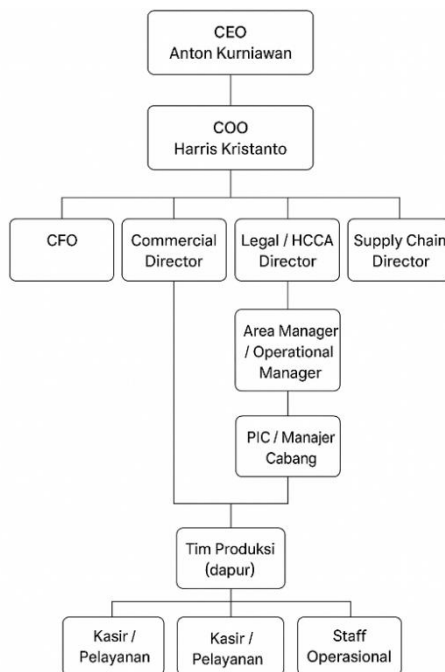
Mie Gacoan Rantauprapat merupakan salah satu cabang dari jaringan restoran Mie Gacoan yang dikenal sebagai tempat makan favorit terutama di kalangan anak muda. Restoran ini berlokasi di Jl. SM. Raja, Jl. Aek Tapa No.7, Rantau Prapat, Kec. Rantau Selatan, Kab. Labuhanbatu, Sumatera Utara. Mie Gacoan menawarkan konsep penyajian cepat dengan menu utama berbasis mie pedas yang memiliki berbagai level kepedasan, serta berbagai pilihan minuman dan camilan pendamping. Produk yang dijual di Mie Gacoan meliputi variasi mie pedas seperti Mie Gacoan dan Mie Hompimpa dengan level kepedasan dari 1 hingga 8, berbagai menu dimsum seperti udang keju, udang rambutan, siomay, dan lumpia udang, serta minuman khas seperti Es Gobak Sodor yang menjadi favorit pelanggan. Popularitasnya tidak hanya berasal dari cita rasa yang khas, tetapi juga dari harga yang terjangkau serta suasana tempat yang dibuat nyaman dan menarik bagi pelanggan.

Sebagai restoran dengan jumlah pengunjung yang tinggi setiap harinya, Mie Gacoan Rantauprapat memiliki tantangan tersendiri dalam mempertahankan kualitas layanan. Tingginya tingkat kunjungan membuat variasi penilaian pelanggan semakin beragam, terutama terkait kecepatan pelayanan, kenyamanan tempat, rasa makanan, serta efisiensi dalam pengelolaan antrean. Oleh karena itu, memahami tingkat kepuasan pelanggan menjadi hal penting untuk menjaga stabilitas bisnis dan meningkatkan kualitas layanan di masa mendatang. Pengelolaan kepuasan pelanggan yang baik juga dapat membantu pihak manajemen dalam menjaga loyalitas serta memperkuat citra merek di mata konsumen.

Dengan banyaknya interaksi dan aktivitas pelayanan, Mie Gacoan Rantauprapat menghasilkan data yang cukup besar dan berpotensi untuk dianalisis lebih lanjut. Data tersebut dapat

memberikan gambaran mengenai pola kepuasan pelanggan jika diolah dengan metode yang tepat. Kondisi ini menjadikan restoran tersebut sebagai objek penelitian yang relevan untuk penerapan teknik Data Mining, khususnya dalam membandingkan performa metode Naive Bayes dan K-Nearest Neighbor (KNN) untuk mengklasifikasikan tingkat kepuasan pelanggan secara lebih akurat dan terukur. Melalui analisis ini, diharapkan diperoleh pemahaman yang lebih mendalam mengenai perilaku pelanggan serta metode terbaik untuk mendukung pengambilan keputusan dalam peningkatan kualitas layanan di Mie Gacoan Rantaupratat.

## b. Struktur Organisasi



**Gambar 4.** Struktur Organisasi

Struktur organisasi pada gambar menunjukkan susunan hierarki kepemimpinan dan pembagian tanggung jawab dalam suatu perusahaan yang bergerak di bidang operasional layanan

makanan. Pada tingkat paling atas terdapat Chief Executive Officer (CEO) yang dijabat oleh Anton Kurniawan. CEO merupakan pemimpin tertinggi dalam organisasi yang bertanggung jawab terhadap arah strategis perusahaan, pengambilan keputusan utama, serta pengawasan terhadap seluruh aktivitas operasional dan manajerial perusahaan. Posisi ini memiliki peran penting dalam menentukan visi, misi, serta kebijakan perusahaan agar dapat berkembang secara berkelanjutan dan kompetitif.

Di bawah CEO terdapat Chief Operating Officer (COO) yang dijabat oleh Harris Kristanto. COO memiliki tanggung jawab utama dalam mengelola kegiatan operasional perusahaan sehari-hari agar berjalan sesuai dengan strategi yang telah ditetapkan oleh CEO. Peran COO sangat penting dalam memastikan bahwa seluruh unit kerja dapat bekerja secara efektif, efisien, dan terkoordinasi dengan baik. COO juga menjadi penghubung antara manajemen tingkat atas dengan berbagai direktorat yang berada di bawahnya, sehingga proses pengambilan keputusan dapat dilaksanakan secara sistematis dan terarah.

Selanjutnya, pada tingkat direktorat terdapat beberapa posisi penting yang mendukung kegiatan operasional perusahaan, yaitu Chief Financial Officer (CFO), Commercial Director, Legal/HCCA Director, dan Supply Chain Director. CFO bertanggung jawab dalam pengelolaan keuangan perusahaan, termasuk perencanaan anggaran, pengawasan arus kas, serta analisis keuangan. Commercial Director berfokus pada strategi pemasaran, pengembangan bisnis, serta peningkatan penjualan. Sementara itu, Legal/HCCA Director bertanggung jawab dalam aspek hukum dan kepatuhan perusahaan terhadap regulasi yang berlaku. Supply Chain Director memiliki peran dalam mengelola rantai pasokan, termasuk pengadaan bahan baku, distribusi, serta pengelolaan logistik agar operasional perusahaan berjalan lancar.

Di bawah struktur direktorat, khususnya pada jalur operasional, terdapat Area Manager atau Operational Manager yang bertanggung jawab mengawasi kinerja operasional di beberapa cabang atau wilayah tertentu. Posisi ini memiliki peran strategis dalam memastikan bahwa setiap cabang menjalankan standar operasional perusahaan secara konsisten. Area Manager juga melakukan koordinasi dengan PIC atau Manajer Cabang yang bertugas memimpin kegiatan operasional pada masing-masing cabang. Manajer cabang memiliki tanggung jawab langsung terhadap pengelolaan sumber daya manusia, pelayanan pelanggan, serta pencapaian target operasional di tingkat cabang.

Pada tingkat operasional paling bawah terdapat Tim Produksi (dapur) yang bertugas menyiapkan produk atau makanan sesuai dengan standar kualitas yang telah ditentukan perusahaan. Tim ini kemudian didukung oleh kasir atau bagian pelayanan yang berperan dalam melayani pelanggan, melakukan transaksi pembayaran, serta memastikan proses pelayanan berjalan dengan baik. Selain itu terdapat staff operasional yang membantu berbagai aktivitas operasional harian di restoran atau cabang. Keseluruhan struktur organisasi ini menunjukkan adanya pembagian tugas yang jelas dari tingkat manajemen strategis hingga tingkat operasional, sehingga setiap bagian dapat bekerja secara terkoordinasi untuk mencapai tujuan perusahaan secara efektif.

### **c. Tugas dan Wewenang**

#### **1. CEO**

Chief Executive Officer (CEO) merupakan pimpinan tertinggi dalam struktur organisasi perusahaan yang memiliki tanggung jawab utama dalam menentukan arah dan strategi perusahaan secara keseluruhan. Posisi ini berperan penting dalam merumuskan visi, misi, serta

kebijakan strategis yang menjadi pedoman bagi seluruh aktivitas operasional perusahaan. Dalam menjalankan tugasnya, CEO harus mampu menganalisis kondisi internal dan eksternal perusahaan, mengidentifikasi peluang dan tantangan bisnis, serta menetapkan langkah-langkah strategis yang dapat mendorong pertumbuhan dan keberlanjutan perusahaan dalam jangka panjang.

Selain menetapkan arah strategis perusahaan, CEO juga memiliki wewenang untuk mengawasi dan mengevaluasi kinerja seluruh divisi yang ada di dalam organisasi. Pengawasan ini dilakukan untuk memastikan bahwa setiap unit kerja menjalankan tugasnya sesuai dengan tujuan, kebijakan, serta standar operasional yang telah ditetapkan. Melalui fungsi pengawasan tersebut, CEO dapat memastikan bahwa setiap kegiatan perusahaan berjalan secara efektif, efisien, dan selaras dengan target yang ingin dicapai.

Dengan wewenang yang dimiliki, CEO juga berperan sebagai pengambil keputusan utama dalam berbagai kebijakan penting perusahaan, baik yang berkaitan dengan pengembangan bisnis, pengelolaan sumber daya, maupun peningkatan kualitas layanan. Oleh karena itu, posisi CEO tidak hanya menuntut kemampuan kepemimpinan yang kuat, tetapi juga kemampuan strategis dalam mengelola organisasi secara menyeluruh sehingga seluruh bagian perusahaan dapat bekerja secara terkoordinasi untuk mencapai tujuan perusahaan.

## 2. COO

Chief Operating Officer (COO) merupakan salah satu posisi penting dalam struktur organisasi perusahaan

yang bertanggung jawab terhadap pengelolaan kegiatan operasional sehari-hari. COO memiliki tugas utama untuk memastikan bahwa seluruh proses operasional perusahaan dapat berjalan secara efektif, efisien, dan sesuai dengan strategi yang telah ditetapkan oleh pimpinan perusahaan. Dalam menjalankan tugasnya, COO berperan dalam mengoordinasikan berbagai unit kerja, mengawasi pelaksanaan kegiatan operasional, serta memastikan bahwa setiap bagian organisasi bekerja sesuai dengan standar operasional yang telah ditetapkan.

Selain itu, COO juga memiliki wewenang dalam mengambil berbagai keputusan operasional tingkat tinggi yang berkaitan dengan kelancaran aktivitas perusahaan. Keputusan tersebut dapat mencakup pengelolaan sumber daya, peningkatan efisiensi operasional, hingga penyelesaian permasalahan yang muncul dalam proses kerja sehari-hari. Dengan kewenangan tersebut, COO berperan sebagai penghubung antara kebijakan strategis yang ditetapkan oleh pimpinan perusahaan dengan implementasinya di tingkat operasional.

Melalui peran tersebut, COO diharapkan mampu memastikan bahwa seluruh kegiatan operasional perusahaan berjalan secara terstruktur, terkoordinasi, dan selaras dengan tujuan organisasi. Oleh karena itu, posisi COO menuntut kemampuan manajerial yang baik, keterampilan koordinasi yang tinggi, serta kemampuan dalam mengambil keputusan yang tepat guna mendukung keberhasilan operasional perusahaan secara keseluruhan.

### 3. CFO

Chief Financial Officer (CFO) merupakan pimpinan yang bertanggung jawab dalam pengelolaan dan pengawasan seluruh aspek keuangan perusahaan. Tugas utama CFO adalah memastikan bahwa kondisi keuangan perusahaan tetap sehat dan stabil melalui perencanaan, pengelolaan, serta pengendalian keuangan yang efektif. Dalam menjalankan perannya, CFO mengelola berbagai aktivitas keuangan seperti penyusunan anggaran perusahaan, pengawasan arus kas, serta penyusunan dan analisis laporan keuangan yang digunakan sebagai dasar dalam pengambilan keputusan manajerial.

Selain itu, CFO juga memiliki tanggung jawab untuk memastikan bahwa pengelolaan keuangan perusahaan dilakukan secara transparan, akuntabel, dan sesuai dengan standar akuntansi yang berlaku. Melalui laporan keuangan yang disusun secara sistematis, CFO dapat memberikan informasi yang akurat kepada pimpinan perusahaan mengenai kondisi finansial perusahaan, sehingga dapat digunakan untuk merencanakan strategi bisnis yang lebih efektif dan berkelanjutan.

Dalam hal wewenang, CFO memiliki hak untuk menyetujui berbagai pengeluaran serta menentukan alokasi dana perusahaan sesuai dengan kebutuhan operasional dan strategi bisnis yang telah ditetapkan. Wewenang ini memungkinkan CFO untuk mengendalikan penggunaan sumber daya keuangan secara optimal serta memastikan bahwa setiap pengeluaran perusahaan dilakukan secara tepat, efisien, dan memberikan manfaat maksimal bagi perkembangan perusahaan.

4. Commercial Director

Commercial Director merupakan pimpinan yang bertanggung jawab dalam mengembangkan dan mengelola strategi pemasaran serta penjualan perusahaan. Tugas utama posisi ini adalah merancang berbagai strategi yang dapat meningkatkan daya saing produk atau layanan perusahaan di pasar. Dalam menjalankan perannya, Commercial Director juga berfokus pada upaya meningkatkan jumlah pelanggan, memperluas pangsa pasar, serta memastikan bahwa target penjualan perusahaan dapat tercapai secara optimal.

Selain itu, Commercial Director memiliki wewenang untuk menentukan kebijakan yang berkaitan dengan harga produk dan strategi promosi yang akan diterapkan. Kebijakan tersebut bertujuan untuk menarik minat konsumen sekaligus menjaga keseimbangan antara keuntungan perusahaan dan daya beli pelanggan. Dengan kewenangan tersebut, Commercial Director berperan penting dalam mendukung pertumbuhan bisnis dan meningkatkan kinerja penjualan perusahaan..

5. Legal / HCCA Director

Legal/HCCA Director merupakan pimpinan yang bertanggung jawab dalam mengawasi aspek hukum serta kepatuhan perusahaan terhadap peraturan dan perundang-undangan yang berlaku. Tugas utama posisi ini adalah memastikan bahwa seluruh kegiatan operasional perusahaan berjalan sesuai dengan ketentuan hukum, sehingga perusahaan dapat terhindar dari risiko hukum yang dapat merugikan organisasi. Selain itu, Legal/HCCA Director juga berperan dalam penyusunan, peninjauan, serta pengelolaan berbagai

dokumen hukum dan perjanjian yang berkaitan dengan kegiatan perusahaan.

Dalam menjalankan tugasnya, Legal/HCCA Director memiliki wewenang untuk memberikan persetujuan terhadap berbagai dokumen hukum yang digunakan oleh perusahaan, seperti kontrak kerja sama, perjanjian bisnis, maupun dokumen legal lainnya. Wewenang ini memungkinkan Legal/HCCA Director untuk memastikan bahwa setiap dokumen yang digunakan telah memenuhi ketentuan hukum yang berlaku serta melindungi kepentingan perusahaan secara menyeluruh. Dengan demikian, peran Legal/HCCA Director sangat penting dalam menjaga kepastian hukum dan keberlangsungan operasional perusahaan.

#### 6. Supply Chain Director

Supply Chain Director merupakan pimpinan yang bertanggung jawab dalam mengelola seluruh aktivitas rantai pasokan perusahaan, khususnya dalam pengadaan bahan baku serta distribusi barang ke setiap cabang. Tugas utama posisi ini adalah memastikan bahwa ketersediaan bahan baku selalu terjaga sehingga proses operasional perusahaan dapat berjalan dengan lancar. Selain itu, Supply Chain Director juga mengatur sistem distribusi agar pengiriman bahan atau produk ke berbagai cabang dapat dilakukan secara tepat waktu, efisien, dan sesuai dengan kebutuhan operasional.

Dalam menjalankan tugasnya, Supply Chain Director memiliki wewenang untuk memilih pemasok atau supplier yang akan bekerja sama dengan perusahaan serta menetapkan standar kualitas bahan baku yang digunakan. Wewenang ini bertujuan untuk memastikan bahwa bahan baku yang diperoleh memiliki

kualitas yang baik dan memenuhi standar perusahaan. Dengan demikian, peran Supply Chain Director sangat penting dalam menjaga kualitas produk sekaligus mendukung kelancaran proses operasional perusahaan.

7. Area Manager / Operational Manager

Area Manager atau Operational Manager merupakan posisi yang bertanggung jawab dalam mengawasi dan mengoordinasikan kegiatan operasional di beberapa cabang dalam suatu wilayah tertentu. Tugas utama posisi ini adalah memastikan bahwa setiap cabang menjalankan kegiatan operasional sesuai dengan standar perusahaan, baik dari segi pelayanan, kualitas produk, maupun kinerja karyawan. Dalam menjalankan perannya, Area Manager juga melakukan pemantauan terhadap pencapaian target operasional serta memberikan arahan kepada manajer cabang agar kegiatan operasional dapat berjalan secara efektif. Selain itu, Area Manager memiliki wewenang untuk mengambil keputusan operasional pada tingkat wilayah atau cabang, terutama dalam menangani permasalahan yang berkaitan dengan operasional harian, sehingga setiap cabang dapat tetap beroperasi dengan baik dan selaras dengan kebijakan perusahaan.

8. PIC / Manajer Cabang

PIC atau Manajer Cabang merupakan pihak yang bertanggung jawab dalam mengelola seluruh kegiatan operasional harian di tingkat cabang. Tugas utama posisi ini adalah memastikan bahwa setiap aktivitas operasional, mulai dari pelayanan pelanggan, pengelolaan staf, hingga pengawasan proses kerja di cabang, dapat berjalan dengan baik dan sesuai dengan standar operasional perusahaan. Manajer Cabang juga

berperan dalam menjaga kualitas pelayanan serta memastikan bahwa target operasional cabang dapat tercapai.

Dalam menjalankan tugasnya, PIC atau Manajer Cabang memiliki wewenang untuk mengatur jadwal kerja staf serta melakukan pembagian tugas kepada setiap karyawan sesuai dengan tanggung jawab masing-masing. Wewenang ini bertujuan agar kegiatan operasional di cabang dapat berjalan secara terorganisir dan efisien. Dengan pengelolaan jadwal dan pembagian tugas yang baik, kinerja tim dapat lebih optimal sehingga pelayanan kepada pelanggan dapat diberikan secara maksimal.

9. Tim Produksi (Dapur)

Tim Produksi (Dapur) merupakan bagian operasional yang bertanggung jawab dalam menyiapkan dan mengolah makanan sesuai dengan standar resep serta prosedur yang telah ditetapkan oleh perusahaan. Tim ini memastikan bahwa setiap produk yang dihasilkan memiliki kualitas, rasa, dan tampilan yang konsisten sehingga dapat memenuhi kepuasan pelanggan. Selain itu, tim produksi juga memiliki wewenang untuk menentukan kebutuhan bahan baku harian yang diperlukan dalam proses pengolahan makanan agar kegiatan operasional dapur dapat berjalan dengan lancar.

10. Kasir / Pelayanan

Kasir/Pelayanan bertugas melakukan proses transaksi pembayaran dengan pelanggan serta memastikan proses pemesanan berjalan dengan baik. Selain itu, bagian ini memiliki wewenang untuk memberikan pelayanan langsung kepada pelanggan, seperti menerima pesanan dan menyajikan makanan.

#### 11. Staff Operasional

Staff Operasional bertugas membantu berbagai kegiatan operasional di cabang, seperti menjaga kebersihan area kerja, melakukan penyetokan barang, serta membantu proses pelayanan ketika dibutuhkan. Peran ini penting untuk memastikan kegiatan operasional sehari-hari dapat berjalan dengan lancar.

Dalam menjalankan tugasnya, Staff Operasional memiliki wewenang untuk mendukung kelancaran aktivitas di berbagai divisi sesuai dengan kebutuhan operasional. Dukungan tersebut membantu menjaga efisiensi kerja serta memastikan pelayanan kepada pelanggan tetap optimal.

### 3. Hasil dan Pembahasan

Hasil pada penelitian ini membahas tentang penerapan metode Naive Bayes dan K-Nearest Neighbor (KNN) dengan memanfaatkan aplikasi RapidMiner sebagai alat bantu dalam proses analisis data. Penggunaan RapidMiner bertujuan untuk mempermudah proses pengolahan data serta implementasi kedua algoritma klasifikasi tersebut secara sistematis. Melalui aplikasi ini, proses analisis dapat dilakukan dengan lebih terstruktur sehingga data yang digunakan dalam penelitian dapat diolah secara optimal.

Proses pengolahan data diawali dengan memasukkan data kepuasan pelanggan ke dalam sistem RapidMiner. Setelah data dimasukkan, dilakukan beberapa tahapan pengolahan data yang meliputi preprocessing, transformasi data, serta pembagian data menjadi dua bagian utama, yaitu data training dan data testing. Tahapan ini dilakukan untuk memastikan bahwa data yang digunakan dalam proses analisis telah berada dalam kondisi yang sesuai sehingga dapat menghasilkan model klasifikasi yang lebih akurat.

Selanjutnya, algoritma Naive Bayes dan K-Nearest Neighbor (KNN) diterapkan pada data yang telah diproses sebelumnya. Kedua algoritma tersebut digunakan untuk menghasilkan keluaran berupa klasifikasi tingkat kepuasan pelanggan berdasarkan beberapa atribut yang digunakan dalam penelitian, yaitu kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Melalui penerapan kedua metode tersebut, sistem dapat menghasilkan prediksi klasifikasi yang menunjukkan kategori kepuasan pelanggan berdasarkan data yang dianalisis.

Hasil yang diperoleh dalam penelitian ini berupa data prediksi klasifikasi serta nilai evaluasi yang dihasilkan dari proses perhitungan masing-masing metode. Dengan demikian, bagian hasil penelitian ini berfokus pada penyajian keluaran dari proses penerapan kedua metode pada RapidMiner tanpa melakukan penarikan kesimpulan mengenai tingkat keunggulan salah satu metode. Hal tersebut disebabkan karena analisis lebih lanjut terhadap hasil yang diperoleh masih akan dilakukan secara lebih mendalam pada tahap pembahasan penelitian.

a. Pengumpulan data

Pengumpulan data dalam penelitian ini dilakukan dengan cara menyebarkan kuesioner secara langsung kepada pelanggan Mie Gacoan Rantauprapat yang sedang melakukan kunjungan dan menikmati layanan di lokasi penelitian. Data yang diperoleh berasal dari jawaban responden terhadap beberapa pertanyaan yang berkaitan dengan tingkat kepuasan pelanggan, khususnya pada aspek kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Setiap responden diminta memberikan penilaian sesuai dengan pengalaman yang mereka rasakan selama berkunjung, sehingga data yang terkumpul dapat mencerminkan kondisi pelayanan yang sebenarnya di lapangan.

Seluruh data hasil kuesioner kemudian direkapitulasi dan dicatat dalam bentuk tabel untuk memudahkan proses pengolahan dan analisis data selanjutnya. Dengan demikian, sumber data dalam penelitian ini berasal langsung dari pelanggan sebagai responden utama, sehingga informasi yang diperoleh bersifat primer. Data tersebut kemudian digunakan sebagai dasar dalam proses analisis dan penerapan metode klasifikasi menggunakan algoritma Naive Bayes dan K-Nearest Neighbor (KNN).

Nama Lengkap	Gmail	Kualitas Pelayanan	Kualitas Produk	Kecepatan Penyajian	Kenyamanan Tempat	Kepuasan
AWS	ax@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
ADL	ax@gmail.com	Baik	Bagus	Lambat	Nyaman	Puas
A	ax@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
Dst.						

**Tabel 1.** Hasil Kuisisioner

Pada tabel tersebut ditampilkan data hasil kuesioner kepuasan pelanggan yang telah dilengkapi dengan atribut Kepuasan sebagai label kelas, yaitu Puas dan Tidak Puas. Data ini memuat identitas responden yang telah disamarkan, alamat Gmail, serta penilaian terhadap empat variabel utama, yaitu kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Setiap responden memberikan penilaian berdasarkan pengalaman mereka saat berkunjung, sehingga data yang dihasilkan dapat menggambarkan kondisi pelayanan yang dirasakan pelanggan. Variasi nilai seperti Baik, Kurang Baik, Bagus, Kurang Bagus, Cepat, Lambat, Nyaman, dan Kurang Nyaman menunjukkan adanya perbedaan persepsi pelanggan terhadap layanan yang diterima.

Data pada tabel tersebut selanjutnya digunakan sebagai dataset dalam penerapan metode klasifikasi Naive Bayes dan K-Nearest Neighbor (KNN). Setiap baris data mewakili satu responden dengan kombinasi nilai atribut yang berbeda, sehingga dapat digunakan sebagai bahan pembelajaran bagi model untuk mengenali pola kepuasan pelanggan. Dataset ini juga menjadi dasar dalam proses pembagian data training dan data testing sebelum dilakukan proses perhitungan menggunakan algoritma klasifikasi, sehingga analisis dapat dilakukan secara lebih sistematis dan objektif.

b. Pembagian Data

Pembagian data dalam penelitian ini dilakukan dengan memisahkan dataset menjadi dua bagian utama, yaitu data training dan data testing, untuk mendukung proses analisis menggunakan metode klasifikasi. Data training digunakan sebagai bahan pembelajaran model dalam mengenali pola hubungan antara atribut kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat terhadap tingkat kepuasan pelanggan, sedangkan data testing digunakan sebagai data uji untuk mengevaluasi kemampuan model dalam melakukan prediksi pada data baru. Pemisahan data ini bertujuan agar model tidak hanya menghafal data yang ada, tetapi juga mampu melakukan generalisasi terhadap data yang belum pernah diproses sebelumnya, sehingga hasil klasifikasi dapat diukur secara lebih objektif dan sistematis.

<b>Nama Lengkap</b>	<b>Gmail</b>	<b>Kualitas Pelayanan</b>	<b>Kualitas Produk</b>	<b>Kecepatan Penyajian</b>	<b>Kenyamanan Tempat</b>	<b>Kepuasan</b>
AWS	axx@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
ADL	axx@gmail.com	Baik	Bagus	Lambat	Nyaman	Puas
A	axx@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
Dst.						

**Tabel 2.** Data Training

<b>Nama Lengkap</b>	<b>Gmail</b>	<b>Kualitas Pelayanan</b>	<b>Kualitas Produk</b>	<b>Kecepatan Penyajian</b>	<b>Kenyamanan Tempat</b>	<b>Kepuasan</b>
AWS	axx@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
ADL	axx@gmail.com	Baik	Bagus	Lambat	Nyaman	Puas
A	axx@gmail.com	Baik	Bagus	Cepat	Nyaman	Puas
Dst.						

**Tabel 3.** Data Testing

Tabel Data Training dan Data Testing menyajikan data responden yang diperoleh dari hasil kuesioner kepuasan pelanggan yang digunakan dalam proses klasifikasi menggunakan metode Naive Bayes dan K-Nearest Neighbor (KNN). Setiap tabel memuat atribut utama seperti nama responden yang telah disamarkan, alamat Gmail, kualitas pelayanan, kualitas produk, kecepatan penyajian, kenyamanan tempat, serta kategori kepuasan pelanggan yang terdiri dari kelas Puas dan Tidak Puas. Data training digunakan sebagai data latih untuk membangun model klasifikasi, sedangkan data testing digunakan sebagai data uji untuk melihat kemampuan model dalam melakukan prediksi.

Kesamaan format antara tabel data training dan data testing menunjukkan bahwa kedua dataset berasal dari sumber data yang sama, yaitu hasil kuesioner pelanggan, namun memiliki fungsi yang berbeda dalam proses penelitian. Data training digunakan untuk mempelajari pola kepuasan pelanggan berdasarkan atribut yang ada, sedangkan data testing digunakan untuk mengevaluasi kinerja model dalam memprediksi tingkat kepuasan pelanggan secara objektif. Pembagian ini membantu memastikan bahwa hasil klasifikasi dapat dianalisis secara lebih sistematis dan dapat dipertanggungjawabkan.

c. Langkah-Langkah Pengolahan Data

Langkah pengolahan data pada penelitian ini dilakukan menggunakan aplikasi RapidMiner melalui beberapa tahapan analisis. Proses dimulai dengan mengimpor dataset hasil kuesioner ke dalam RapidMiner, kemudian dilakukan data cleaning untuk memastikan tidak terdapat data kosong, duplikasi, maupun kesalahan pengisian yang dapat memengaruhi hasil analisis. Selanjutnya dilakukan transformasi data dengan menyesuaikan data kategorikal agar dapat diproses oleh algoritma klasifikasi.

Setelah proses tersebut, atribut yang digunakan sebagai variabel input ditentukan, sedangkan atribut kepuasan dijadikan sebagai label atau kelas tujuan. Data yang telah dipersiapkan kemudian diproses menggunakan algoritma Naive Bayes dan K-Nearest Neighbor (KNN) untuk menghasilkan klasifikasi tingkat kepuasan pelanggan berdasarkan atribut kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Hasil pemodelan selanjutnya dievaluasi menggunakan fitur evaluasi pada RapidMiner untuk memperoleh informasi mengenai kinerja model klasifikasi.

## 1. Buka Aplikasi RapidMiner

Langkah pertama dalam proses pengolahan data adalah membuka aplikasi RapidMiner Studio yang telah terinstal pada perangkat. Setelah aplikasi dijalankan, pengguna akan masuk ke halaman utama yang menampilkan workspace sebagai area kerja untuk melakukan analisis data. Pada tampilan ini tersedia beberapa menu utama seperti Repository, Operators, Process, Parameters, dan Help yang berfungsi untuk mendukung proses pengolahan data.

Panel Repository digunakan sebagai tempat penyimpanan dataset penelitian, sedangkan panel Operators berisi berbagai alat untuk melakukan preprocessing data, pemodelan, dan evaluasi. Area Process berfungsi sebagai tempat menyusun alur kerja analisis data secara sistematis, sementara panel Parameters digunakan untuk mengatur konfigurasi operator yang digunakan dalam proses analisis.

## 2. Blank Process

Blank Process merupakan lembar kerja kosong pada RapidMiner yang digunakan untuk menyusun alur proses pengolahan data. Pada halaman ini, pengguna dapat menambahkan berbagai operator seperti input data, preprocessing, pemodelan, dan evaluasi untuk membangun proses data mining sesuai kebutuhan penelitian.

Tampilan Blank Process memberikan fleksibilitas bagi peneliti dalam merancang workflow analisis data secara sistematis. Dengan tampilan visual yang tersedia, peneliti dapat memantau setiap tahapan proses analisis serta menghubungkan operator yang diperlukan untuk membentuk alur pengolahan data yang terstruktur.

### 3. Tambahkan Operators

Menambahkan operators pada RapidMiner merupakan langkah penting dalam membangun alur proses analisis data. Operator berfungsi sebagai alat untuk menjalankan berbagai tahapan seperti memasukkan dataset, melakukan preprocessing, membangun model klasifikasi, hingga mengevaluasi hasil analisis.

Salah satu operator yang digunakan adalah Read Excel, yang berfungsi untuk mengimpor dataset dari file Microsoft Excel ke dalam RapidMiner. Setelah operator ini ditambahkan ke area proses, dataset dapat dibaca oleh sistem dan selanjutnya diproses pada tahap analisis berikutnya seperti preprocessing, pemodelan, dan evaluasi.

### 4. Import Data

Tahap import data bertujuan untuk memasukkan dataset ke dalam sistem RapidMiner agar dapat dianalisis lebih lanjut. Dataset yang berasal dari file Excel atau CSV diunggah ke dalam workspace sehingga dapat dikenali sebagai objek data oleh sistem.

Setelah data berhasil diimpor, RapidMiner menampilkan seluruh atribut dan record responden dalam bentuk tabel. Tampilan ini memudahkan peneliti untuk melakukan pengecekan awal terhadap struktur data, seperti memastikan tidak terdapat kesalahan format atau data yang tidak lengkap sebelum melanjutkan ke tahap analisis berikutnya.

### 5. Hubungkan Operators

Menghubungkan seluruh operator merupakan langkah penting untuk membentuk alur pengolahan data yang sistematis. Setiap operator dihubungkan sesuai tahapan penelitian, mulai dari import data, preprocessing, pemodelan, hingga evaluasi hasil, sehingga proses analisis dapat berjalan secara terintegrasi.

Pada penelitian ini, operator Read Excel digunakan untuk mengimpor dataset, kemudian Set Role digunakan untuk menentukan atribut kepuasan sebagai label klasifikasi. Selanjutnya digunakan operator Naive Bayes dan KNN untuk membangun model klasifikasi. Setelah model terbentuk, operator Apply Model digunakan untuk menghasilkan prediksi, dan tahap akhir dilakukan menggunakan operator Performance untuk mengevaluasi kinerja model.

#### 6. Hasil Klasifikasi

Hasil klasifikasi merupakan keluaran dari proses pengolahan data menggunakan metode Naive Bayes dan K-Nearest Neighbor (KNN) pada aplikasi RapidMiner. Proses ini menghasilkan pengelompokan data responden ke dalam dua kategori, yaitu Puas dan Tidak Puas, berdasarkan atribut kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat.

Dari hasil klasifikasi yang diperoleh, terdapat 27 responden masuk kategori Puas dan 23 responden masuk kategori Tidak Puas. Hasil ini menunjukkan adanya variasi tingkat kepuasan pelanggan berdasarkan penilaian terhadap atribut layanan yang dianalisis.

#### 7. Hasil Performance

Hasil performance merupakan tahap evaluasi untuk mengetahui kemampuan model dalam melakukan klasifikasi data. RapidMiner menghitung beberapa indikator kinerja seperti akurasi, precision, dan recall berdasarkan perbandingan antara label asli data dengan hasil prediksi model.

Hasil evaluasi menunjukkan bahwa model klasifikasi memiliki tingkat akurasi sebesar 100%, dengan nilai precision dan recall juga mencapai 100% untuk kedua kelas. Hal ini menunjukkan bahwa model mampu mengenali pola kepuasan pelanggan dengan sangat baik berdasarkan atribut yang digunakan dalam penelitian.

## **4. Penutup**

### **a. Kesimpulan**

Penerapan metode Naive Bayes dan K-Nearest Neighbor (KNN) dapat digunakan untuk mengklasifikasikan tingkat kepuasan pelanggan di Mie Gacoan Rantauprapat berdasarkan variabel kualitas pelayanan, kualitas produk, kecepatan penyajian, dan kenyamanan tempat. Proses klasifikasi dilakukan melalui tahapan pengumpulan data, pembagian data menjadi data training dan data testing, transformasi data, serta perhitungan menggunakan kedua metode tersebut.

Teknik Data Mining terbukti mampu membantu mengolah data kepuasan pelanggan menjadi informasi yang lebih terstruktur dan sistematis. Melalui proses klasifikasi, pola kepuasan pelanggan dapat diidentifikasi sehingga memberikan gambaran yang lebih jelas mengenai kondisi kepuasan pelanggan berdasarkan data yang diperoleh.

Hasil penelitian ini menunjukkan bahwa analisis menggunakan metode Naive Bayes dan KNN dapat digunakan sebagai dasar dalam mendukung pengambilan keputusan untuk meningkatkan kualitas layanan. Dengan mengetahui faktor-faktor yang memengaruhi kepuasan pelanggan, pihak Mie Gacoan Rantauprapat dapat melakukan evaluasi serta perbaikan terhadap aspek pelayanan yang masih perlu ditingkatkan.

## **b. Saran**

Berdasarkan hasil penelitian yang telah dilakukan, disarankan agar pihak Mie Gacoan Rantauprapat terus meningkatkan kualitas pelayanan, kualitas produk, kecepatan penyajian, serta kenyamanan tempat karena faktor-faktor tersebut terbukti berpengaruh terhadap tingkat kepuasan pelanggan. Peningkatan kualitas layanan secara berkelanjutan dapat membantu mempertahankan loyalitas pelanggan serta meningkatkan citra positif restoran di mata masyarakat. Selain itu, pengumpulan data kepuasan pelanggan secara rutin juga perlu dilakukan agar pihak manajemen dapat mengetahui perubahan persepsi pelanggan dari waktu ke waktu.

Bagi peneliti selanjutnya, disarankan untuk menggunakan jumlah data yang lebih banyak serta menambahkan variabel lain yang dapat memengaruhi kepuasan pelanggan agar hasil analisis menjadi lebih akurat. Selain itu, penelitian selanjutnya juga dapat mencoba membandingkan metode Data Mining lainnya untuk mengetahui metode klasifikasi yang paling efektif dalam menganalisis kepuasan pelanggan pada sektor kuliner maupun bidang lainnya. Dengan demikian, hasil penelitian dapat memberikan kontribusi yang lebih luas dalam pengembangan analisis data menggunakan teknik Data Mining.



# DAFTAR PUSTAKA





## DAFTAR PUSTAKA

- Abas, M. I., Ibrahim, I., Syahrial, S., Lamusu, R., Baderan, U. S., & Kango, R. (2023). Analysis of Covid-19 Growth Trends Through Data Mining Approach As Decision Support. *Sinkron*, 8(1), 101–108. <https://doi.org/10.33395/sinkron.v8i1.11861>
- Adjani, K., Fauzia, F. A., & Juliane, C. (2023). Comparison of K-Nearest Neighbor and Naïve Bayes Algorithms for Prediction of Aptikom Membership Activity Extension in 2023. *Sinkron*, 8(2), 700–707. <https://doi.org/10.33395/sinkron.v8i2.12081>
- Alam, A., Alana, D. A. F., & Juliane, C. (2023). Comparison Of The C.45 And Naive Bayes Algorithms To Predict Diabetes. *Sinkron*, 8(4), 2641–2650. <https://doi.org/10.33395/sinkron.v8i4.12998>
- Apriyani, M. E., Maskuri, R. A., Ratsanjani, M. H., Pramudhita, A. N., & Rawansyah, R. (2023). Digital Forensic Investigates Sexual Harassment on Telegram using Naïve Bayes. *Sinkron*, 8(3), 1409–1417. <https://doi.org/10.33395/sinkron.v8i3.12514>
- Arifuddin, N. A., Pinastawa, I. W. R., Anugraha, N., & Pradana, M. G. (2023). Classification of Stroke Opportunities with Neural Network and K-Nearest Neighbor Approaches. *Sinkron*, 8(2), 688–693. <https://doi.org/10.33395/sinkron.v8i2.12228>
- Bustomi, Y., Nugraha, A., Juliane, C., & Rahayu, S. (2023). Data Mining Selection of Prospective Government Employees with Employment Agreements using Naive Bayes Classifier. *Sinkron*, 8(1), 1–8. <https://doi.org/10.33395/sinkron.v8i1.11968>
- Harahap, F., Fahrozi, W., Adawiyah, R., Siregar, E. T., & Harahap, A. Y. N. (2023). Implementasi Data Mining dalam Memprediksi Produk AC Terlaris untuk Meningkatkan Penjualan Menggunakan Metode Naive Bayes. *Jurnal Unitek*, 16(1), 41–51. <https://doi.org/10.52072/unitek.v16i1.541>

- Hasibuan, F. F., Dar, M. H., & Yanris, G. J. (2023). Implementation of the Naïve Bayes Method to determine the Level of Consumer Satisfaction. *Sinkron*, 8(2), 1000–1011. <https://doi.org/10.33395/sinkron.v8i2.12349>
- Hasibuan, S. A., Sihombing, V., & Nasution, F. A. (2023). Analysis of Community Satisfaction Levels using the Neural Network Method in Data Mining. *Sinkron*, 8(3), 1724–1735. <https://doi.org/10.33395/sinkron.v8i3.12634>
- Hermawan, F., & Prianggono, J. (2023). Crime of theft prediction using Machine Learning K-Nearest Neighbour Algorithm at Polresta Bandar Lampung. *Sinkron*, 8(3), 1515–1527. <https://doi.org/10.33395/sinkron.v8i3.12422>
- Indah, I. C., Sari, M. N., & Dar, M. H. (2023). Application of the K-Means Clustering Algorithm to Group Train Passengers in Labuhanbatu. *Sinkron*, 8(2), 825–837. <https://doi.org/10.33395/sinkron.v8i2.12260>
- Irmayani, D., Sinaga, F. A., & Masrizal, M. (2023). Analysis of the Level of Public Satisfaction on the Tiktok Application as an E-Commerce. *Sinkron*, 8(4), 2579–2591. <https://doi.org/10.33395/sinkron.v8i4.13040>
- Kharis, S. A. A., Zili, A. H. A., Putri, A., & Robiansyah, A. (2023). Analisis Tren Minat Masyarakat Indonesia terhadap Artificial Intelligence dalam Menyongsong Society 5.0: Studi Menggunakan Google Trends. *G-Tech: Jurnal Teknologi Terapan*, 7(4), 1345–1354. <https://doi.org/10.33379/gtech.v7i4.3091>
- Kurniadi Hermawan, A., Nugroho, A., & Edora. (2023). Analisa Data Mining Untuk Prediksi Penyakit Ginjal Kronik Dengan Algoritma Regresi Linier. *Bulletin of Information Technology (BIT)*, 4(1), 37–48. <https://doi.org/10.47065/bit.v4i1.475>
- Nasution, R. F., Dar, M. H., & Nasution, F. A. (2023). Implementation of the Naïve Bayes Method to Determine

- Student Interest in Gaming Laptops. *Sinkron*, 8(3), 1709–1723.  
<https://doi.org/10.33395/sinkron.v8i3.12562>
- Punkastyo, D. A., Septian, F., & Syaripudin, A. (2024). Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Prediksi Kelulusan Siswa. *Journal of System and Computer Engineering (JSCE)*, 5(1), 24–35.  
<https://doi.org/10.61628/jsce.v5i1.1073>
- Purwati, N., Pedliyansah, Y., Kurniawan, H., Karnila, S., & Herwanto, R. (2023). Komparasi Metode Apriori dan FP-Growth Data Mining Untuk Mengetahui Pola Penjualan. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(2), 155–161.  
<https://doi.org/10.30591/jpit.v8i2.4876>
- Rahman, R., & Fauzi Abdulloh, F. (2023). Performance of Various Naïve Bayes Using GridSearch Approach In Phishing Email Dataset. *Sinkron*, 8(4), 2336–2344.  
<https://doi.org/10.33395/sinkron.v8i4.12958>
- Rozi, M. F. (2023). Penerapan Data Mining Menggunakan Metode Naive Bayes Untuk Klasifikasi Data Penentuan Hasil Penjualan Dalam Strategi Pemasaran. *Jurnal Komputer Teknologi Informasi Dan Sistem Informasi (JUKTISI)*, 2(2), 444–454.  
<https://doi.org/10.62712/juktisi.v2i2.137>
- Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine. *Sinkron*, 8(3), 1366–1378. <https://doi.org/10.33395/sinkron.v8i3.12447>
- Sari, M., Yanris, G. J., & Hasibuan, M. N. S. (2023). Analysis of the Neural Network Method to Determine Interest in Buying Pertamina Fuel. *Sinkron*, 8(2), 1031–1039.  
<https://doi.org/10.33395/sinkron.v8i2.12292>
- Siregar, A. P., Irmayani, D., & Sari, M. N. (2023). Analysis of the Naïve Bayes Method for Determining Social Assistance

- Eligibility Public. Sinkron, 8(2), 805–817.  
<https://doi.org/10.33395/sinkron.v8i2.12259>
- Supendar, H., Rusdiansyah, R., Suharyanti, N., & Tuslaela, T. (2023). Application of the Naïve Bayes Algorithm in Determining Sales Of The Month. Sinkron, 8(2), 873–879.  
<https://doi.org/10.33395/sinkron.v8i2.12293>
- Zai, F., Sirait, J., Nainggolan, D. W., Sihombing, N. G. D., & Banjarnahor, J. (2023). Comparison Analysis of C4.5 Algorithm and KNN Algorithm for Predicting Data of Non-Active Students at Prima Indonesia University. Sinkron, 8(4), 2027–2035. <https://doi.org/10.33395/sinkron.v8i4.12879>