

BAB III

ANALISIS DAN PERANCANGAN

3.1. Arsitektur Sistem

Arsitektur sistem pada penelitian ini dirancang untuk mendukung proses prediksi tingkat kepuasan pelanggan CS Float secara terstruktur dan berbasis data. Sistem diawali dengan proses pengumpulan data pelanggan yang diperoleh melalui kuesioner atau database, kemudian data tersebut disimpan dan dikelola dalam sebuah basis data. Selanjutnya, data diproses melalui tahap preprocessing yang meliputi seleksi atribut, pembersihan data, dan transformasi data agar siap digunakan dalam proses analisis. Data yang telah diproses kemudian dibagi menjadi data latih dan data uji untuk membangun model klasifikasi. Pada tahap pemodelan, metode Naive Bayes dan K-Nearest Neighbor (KNN) diterapkan untuk mengklasifikasikan tingkat kepuasan pelanggan ke dalam kategori puas dan tidak puas. Hasil klasifikasi dari kedua metode tersebut selanjutnya dievaluasi menggunakan metrik akurasi dan performa model untuk menentukan metode yang paling optimal.

3.2. Desain Aktivitas Sistem

Desain aktivitas sistem pada penelitian ini menggambarkan alur kerja sistem dalam memproses data pelanggan hingga menghasilkan prediksi tingkat kepuasan. Aktivitas dimulai dari pengumpulan data pelanggan melalui kuesioner atau sumber data yang tersedia, kemudian data tersebut dimasukkan ke dalam sistem untuk dilakukan proses seleksi dan pembersihan data. Selanjutnya, sistem melakukan

preprocessing data seperti transformasi atribut dan normalisasi nilai agar sesuai dengan kebutuhan algoritma. Setelah itu, data dibagi menjadi data latih dan data uji, lalu diproses menggunakan metode Naive Bayes dan K-Nearest Neighbor (KNN) untuk melakukan klasifikasi tingkat kepuasan pelanggan ke dalam kategori puas dan tidak puas. Hasil klasifikasi dari kedua metode kemudian dievaluasi menggunakan metrik kinerja seperti akurasi untuk mengetahui tingkat keberhasilan sistem. Aktivitas sistem diakhiri dengan penyajian hasil prediksi dan evaluasi model yang dapat digunakan sebagai dasar pengambilan keputusan dalam peningkatan kualitas produk CS Float.

3.3. Penerapan Metode Naïve Bayes

Metode Naive Bayes diterapkan dalam penelitian ini sebagai salah satu teknik klasifikasi untuk memprediksi tingkat kepuasan pelanggan terhadap produk CS Float. Metode ini bekerja berdasarkan prinsip probabilitas dengan menghitung kemungkinan suatu data pelanggan termasuk ke dalam kategori puas atau tidak puas berdasarkan karakteristik yang dimilikinya. Naive Bayes mengasumsikan bahwa setiap atribut seperti rasa, harga, kemasan, dan frekuensi pembelian saling independen, sehingga perhitungan probabilitas dapat dilakukan secara sederhana dan efisien.

3.3.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan menggunakan kuesioner yang dibuat melalui platform Google Form dan disebarakan kepada pelanggan produk minuman CS Float. Kuesioner tersebut dirancang untuk memperoleh data karakteristik pelanggan serta persepsi mereka terhadap produk,

seperti penilaian terhadap rasa, harga, kemasan, dan tingkat kepuasan secara keseluruhan. Penggunaan Google Form dipilih karena memudahkan proses penyebaran kuesioner, pengumpulan respons secara cepat, serta pengelolaan data dalam bentuk digital yang siap diolah untuk keperluan analisis dan pemodelan prediksi tingkat kepuasan pelanggan.

Tabel 3. 1. Data Hasil Kuesioner

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Bintang syukuriyah siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Nurul Aida Siregar	Cepat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Irwansyah, S.T	Cepat	Kurang Bagus	Bagus	Murah	Bersih	Puas
Nur indah nasution	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Aini Minta Ito Harahap	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Eka Fitri r	Cepat	Bagus	Kurang Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Mahal	Bersih	Puas
Diana syahfitri	Cepat	Bagus	Bagus	Murah	Kurang Bersih	Puas
dela yuuliyani	Lambat	Kurang Bagus	Bagus	Mahal	Kurang Bersih	Tidak Puas
Amar ma'rub siregar	Lambat	Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Nurhayati	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Bersih	Tidak Puas
putreimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Grifaril	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Zuhri Dio alfalalah	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Semangat abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
IKA MAYA SARI SIMAMORA	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurul aminah fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Hani aura arcika	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Dinda cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia rambe	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn syafrida	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn	Cepat	Bagus	Bagus	Murah	Bersih	Puas
dhea allya ginting	Cepat	Bagus	Bagus	Murah	Bersih	Puas
SHOLE HAMID HSB	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas merupakan data yang telah diperoleh dari hasil pengumpulan kuesioner yang disebarkan kepada responden melalui Google Form. Data tersebut memuat beberapa atribut penilaian, yaitu waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat, yang digunakan untuk menggambarkan persepsi responden terhadap produk dan layanan CS Float. Setiap responden juga diklasifikasikan ke dalam kategori kepuasan, yaitu puas dan tidak puas, berdasarkan jawaban yang diberikan.

3.3.2. Pembersihan Data

Pembersihan data merupakan tahap awal dalam pengolahan data yang bertujuan untuk memastikan kualitas data sebelum dilakukan analisis lebih lanjut. Pada tahap ini, data hasil kuesioner diperiksa untuk menghilangkan data ganda, memperbaiki kesalahan penulisan, serta menangani data yang tidak lengkap atau

tidak konsisten. Proses pembersihan data dilakukan agar data yang digunakan dalam penelitian bersifat akurat, relevan, dan siap digunakan dalam proses pemodelan, sehingga hasil prediksi tingkat kepuasan pelanggan dapat diperoleh secara optimal.

Tabel 3. 2. Data Setelah Dibersihkan

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Aini Minta Ito Harahap	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Amar Ma'rub Siregar	Lambat	Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Bintang Syukuriyah Siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dela Yuuliyani	Lambat	Kurang Bagus	Bagus	Mahal	Kurang Bersih	Tidak Puas
Dhea Allya Ginting	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Diana Syahfitri	Cepat	Bagus	Bagus	Murah	Kurang Bersih	Puas
Dinda Cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia Rambe	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Eka Fitri R	Cepat	Bagus	Kurang Bagus	Murah	Bersih	Puas
Grifaril	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Hani Aura Arcika	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ika Maya Sari Simamora	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Irwansyah, S.T	Cepat	Kurang Bagus	Bagus	Murah	Bersih	Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nur Indah Nasution	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas

Nurhayati	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Bersih	Tidak Puas
Nurul Aida Siregar	Cepat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Nurul Aminah Fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Putreimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn Syafrida	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Mahal	Bersih	Puas
Semangat Abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Shole Hamid Hsb	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria Syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Zuhri Dio Alfalah	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas merupakan data hasil kuesioner yang telah melalui tahap pembersihan data (data cleaning) sebelum digunakan dalam proses analisis lebih lanjut. Pada tahap ini, data telah diperiksa dan diseleksi sehingga tidak terdapat data ganda, kesalahan penulisan, maupun data yang tidak lengkap. Seluruh atribut seperti waktu respon, kualitas layanan, kualitas produk, harga produk, kebersihan tempat, dan kategori kepuasan telah diseragamkan penulisannya agar konsisten dan mudah diolah.

3.3.3. Pembagian Data

Pada penerapan metode Naive Bayes, data yang telah dibersihkan dibagi menjadi dua bagian, yaitu data training dan data testing. Data training digunakan untuk membentuk model Naive Bayes dengan mempelajari pola hubungan antara atribut penilaian dan kategori kepuasan pelanggan, sedangkan data testing

digunakan untuk menguji kemampuan model Naive Bayes dalam melakukan klasifikasi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data ini bertujuan untuk memperoleh hasil prediksi yang objektif serta mengukur tingkat keakuratan metode Naive Bayes dalam memprediksi kepuasan pelanggan.

Tabel 3. 3. Data Training

Nama Lengkap	Waktu Respon	Kualiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Aini Minta Ito Harahap	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Amar Ma'rub Siregar	Lambat	Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Bintang Syukuriyah Siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Dela Yuuliyani	Lambat	Kurang Bagus	Bagus	Mahal	Kurang Bersih	Tidak Puas
Diana Syahfitri	Cepat	Bagus	Bagus	Murah	Kurang Bersih	Puas
Eka Fitri R	Cepat	Bagus	Kurang Bagus	Murah	Bersih	Puas
Grifaril	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ika Maya Sari Simamora	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Irwansyah, S.T	Cepat	Kurang Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nur Indah Nasution	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Nurhayati	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Bersih	Tidak Puas
Nurul Aida Siregar	Cepat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Putrireimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Mahal	Bersih	Puas
Semangat Abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria Syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Zuhri Dio Alfalah	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas merupakan data training yang digunakan dalam penelitian untuk membangun model klasifikasi tingkat kepuasan pelanggan CS Float. Data ini terdiri

dari beberapa atribut utama, yaitu waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat, yang masing-masing memiliki nilai kategorikal. Setiap kombinasi atribut tersebut diakhiri dengan label kelas Puas atau Tidak Puas sebagai target klasifikasi. Berdasarkan data training tersebut, terlihat bahwa pelanggan dengan waktu respon cepat, kualitas layanan dan produk yang baik, harga terjangkau, serta kebersihan tempat yang baik cenderung dikategorikan puas, sedangkan pelanggan dengan waktu respon lambat, kualitas layanan dan produk yang kurang baik, serta kondisi kebersihan yang kurang cenderung masuk dalam kategori tidak puas.

Tabel 3. 4. Data Testing

Nama Lengkap	Waktu Respon	Kualiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dhea Allya Ginting	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia Rambe	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Hani Aura Arcika	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurul Aminah Fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn Syafrida	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Shole Hamid Hsb	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas menunjukkan pembagian dataset penelitian ke dalam dua bagian, yaitu data training dan data testing, yang digunakan dalam penerapan metode Naive Bayes. Data training berisi data pelanggan yang digunakan untuk membangun dan melatih model klasifikasi agar mampu mengenali pola hubungan antara atribut penilaian dengan kategori kepuasan pelanggan. Sementara itu, data testing digunakan untuk menguji kemampuan model dalam memprediksi tingkat kepuasan pelanggan berdasarkan data yang belum pernah dipelajari sebelumnya. Pembagian data ini dilakukan untuk memastikan bahwa proses pengujian model berjalan secara objektif dan hasil prediksi yang diperoleh dapat menggambarkan kinerja metode Naive Bayes secara akurat.

3.3.4. Transformasi Data

Transformasi data merupakan tahap pengolahan lanjutan yang bertujuan untuk menyesuaikan struktur data training agar mudah dianalisis menggunakan metode Naive Bayes. Pada tahap ini, data training yang semula berada dalam satu tabel utama diubah dengan cara memisahkan setiap variabel ke dalam tabel tersendiri, seperti tabel waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Pemisahan ini dilakukan untuk mempermudah proses perhitungan probabilitas pada masing-masing variabel terhadap kategori kepuasan pelanggan.

Tabel 3. 5. Variabel Waktu Respon

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Waktu Respon	Cepat	13	1	13/14	1/6
	Lambat	1	5	1/14	5/6
	Total	14	6	100%	100%

Tabel 3. 6. Variabel Kualitas Layanan

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Kualitas Layanan	Bagus	13	1	13/14	1/6
	Kurang Bagus	1	5	1/14	5/6
	Total	14	6	100%	100%

Tabel 3. 7. Variabel Kualitas Produk

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Kualitas Produk	Bagus	13	1	13/14	1/6
	Kurang Bagus	1	5	1/14	5/6
	Total	14	6	100%	100%

Tabel 3. 8. Variabel Harga Produk

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Harga Produk	Murah	13	2	13/14	2/6
	Mahal	1	4	1/14	4/6
	Total	14	6	100%	100%

Tabel 3. 9. Variabel Kebersihan Tempat

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Kebersihan Tempat	Bersih	13	1	13/14	1/6
	Kurang Bersih	1	5	1/14	5/6
	Total	14	6	100%	100%

Tabel 3. 10. Variabel Kategori

Kategori		P (Puas) dan P (Tidak Puas)
Puas	14	14/20
Tidak Puas	6	6/20
Total	20	100%

3.3.5. Perhitungan Data

Pada perhitungan metode Naive Bayes terdapat perhitungan yang dilakukan secara manual yaitu sebagai berikut.

Rumus Metode Naive Bayes,
$$P(A | B) = \frac{P(B | A) P(A)}{P(B)} \quad [50]$$

Information:

A : hipotesis data A (kelas tertentu)

B : data dengan kelas yang tidak diketahui

$P(A|B)$: Probabilitas hipotesis berdasarkan kondisi B

$P(A)$: Kemungkinan hipotesis A

$P(B|A)$: Probabilitas B ketika kondisi A

$P(B)$: Probabilitas

Perhitungan

Pada perhitungan yang akan dilakukan, pertama penulis akan menghitung data Adelia Putri. Adapun perhitungannya sebagai berikut.

$$P(\text{Kategori}) = P(\text{Waktu Respon|Cepat}) \times P(\text{Kualitas Layanan|Bagus}) \times P(\text{Kualitas Produk|Bagus}) \times P(\text{Harga Produk|Murah}) \times P(\text{Kebersihan Tempat|Bersih}) \times P(\text{Kategori|Puas})$$

$$P(\text{Puas}) = P(\text{Cepat|Puas}) \times P(\text{Bagus|Puas}) \times P(\text{Bagus|Puas}) \times P(\text{Murah|Puas}) \times P(\text{Bersih|Puas}) \times P(\text{Kategori|Puas})$$

$$= \left(\frac{13}{14}\right) \times \left(\frac{13}{14}\right) \times \left(\frac{13}{14}\right) \times \left(\frac{13}{14}\right) \times \left(\frac{13}{14}\right) \times \left(\frac{20}{40}\right)$$

$$= 0,4869 \text{ (Nilai Puas)}$$

$$P(\text{Tidak Puas}) = P(\text{Cepat|Tidak Puas}) \times P(\text{Bagus|Tidak Puas}) \times P(\text{Bagus|Tidak Puas}) \times P(\text{Murah|Tidak Puas}) \times P(\text{Bersih|Tidak Puas}) \times P(\text{Kategori|Tidak Puas})$$

$$= \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{2}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{20}{40}\right)$$

$$= 0,00006 \text{ (Nilai Tidak Puas)}$$

Berdasarkan hasil perhitungan pada data milik Adelia Putri, diperoleh nilai probabilitas kategori “Puas” sebesar 0,4869, sedangkan nilai probabilitas kategori “Tidak Puas” hanya sebesar 0,00006. Perbandingan ini menunjukkan bahwa

peluang Adelia Putri masuk ke dalam kelas Puas jauh lebih besar dibandingkan peluang masuk kelas Tidak Puas. Dengan demikian, dapat disimpulkan bahwa Adelia Putri termasuk pelanggan yang merasa Puas terhadap CS Float. Hasil perhitungan untuk data pelanggan lainnya akan dipaparkan dalam bentuk tabel pada bagian berikutnya untuk memberikan gambaran yang lebih jelas mengenai distribusi klasifikasi kepuasan pelanggan.

Tabel 3. 11. Data Hasil Perhitungan Metode Naïve Bayes

Nama Lengkap	Kategori	Kategori	Keterangan
Adelia Putri	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Adinda Sevira	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Arnes Dian Putri Harefa	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Dea	Puas	0.483253	Puas
	Tidak Puas	0.000077	
dhea allya ginting	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Dinda cantiks	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Dinda Olivia rambe	Puas	0.000001	Tidak Puas
	Tidak Puas	0.096451	
Divan Alfaruzi	Puas	0.000001	Tidak Puas
	Tidak Puas	0.096451	
Hani aura arcika	Puas	0.000001	Tidak Puas
	Tidak Puas	0.096451	
Husna	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Indah Rahmadani	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Madzy	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Nurul aminah fajri	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Ririn	Puas	0.483253	Puas
	Tidak Puas	0.000077	

Ririn syafrida	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Rizalunsani Rambe	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Shole Hamid Hsb	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Sindi Rambe	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Solehuddin Siregar	Puas	0.483253	Puas
	Tidak Puas	0.000077	
Tika Andini	Puas	0.483253	Puas
	Tidak Puas	0.000077	

Hasil klasifikasi menggunakan metode Naive Bayes menunjukkan bahwa sebagian besar data pelanggan memiliki probabilitas kategori Puas yang jauh lebih tinggi dibandingkan probabilitas Tidak Puas. Hal ini tampak dari nilai probabilitas yang berulang, seperti 0,483253 untuk kategori Puas dan 0,000077 untuk kategori Tidak Puas, yang konsisten pada mayoritas nama dalam tabel. Perbedaan nilai probabilitas ini memberikan indikasi kuat bahwa atribut-atribut yang digunakan dalam proses perhitungan lebih mendukung kecenderungan pelanggan untuk merasa puas terhadap produk CS Float.

Beberapa pelanggan menunjukkan hasil yang berbeda, misalnya Dinda Olivia Rambe, Divan Alfaruzi, dan Hani Aura Arcika, yang memiliki nilai probabilitas Puas sangat rendah (0,000001) dan probabilitas Tidak Puas sebesar 0,096451. Kondisi ini menyebabkan ketiga nama tersebut diklasifikasikan sebagai Tidak Puas, karena peluang untuk masuk kategori tersebut lebih besar dibandingkan dengan kategori Puas. Hal ini sekaligus menunjukkan bahwa metode Naive Bayes mampu mengidentifikasi perbedaan kecenderungan berdasarkan pola data, meskipun rentang probabilitasnya cukup kecil.

Secara keseluruhan, pola perhitungan pada tabel memperlihatkan bahwa kelompok pelanggan yang puas mendominasi hasil klasifikasi, sementara hanya sebagian kecil yang dikategorikan tidak puas. Dominasi ini memperkuat kesimpulan bahwa respon terhadap CS Float cenderung positif berdasarkan data pelanggan yang dianalisis. Dengan demikian, model Naive Bayes tidak hanya memberikan hasil klasifikasi, tetapi juga menggambarkan kecenderungan probabilistik yang dapat dijadikan acuan dalam memahami persepsi pelanggan terhadap kualitas CS Float.

3.3.6. Evaluasi Model

Evaluasi metode Naive Bayes pada penelitian ini menggunakan Confusion Matrix. Hasil tersebut yang nantinya akan memberikan nilai efektivitas dari metode yang digunakan.

		Predicted		Σ
		Puas	Tidak Puas	
Actual	Puas	17	0	17
	Tidak Puas	0	3	3
Σ		17	3	20

Hasil True Positive (TP) adalah 17. True Negative (TN) adalah 23, False Positive (FP) adalah 0 dan False Negative (FN) adalah 0. Maka Nilai akurasi, presisi dan recall adalah sebagai berikut:

$$\text{Accuracy} = \frac{17+3}{17+3+0+0} \times 100\% \quad \text{Then the Accuracy value} = 100\%$$

$$\text{Presisi} = \frac{17}{17+0} \times 100\% \quad \text{Then the Precision value} = 100\%$$

$$\text{Recall} = \frac{17}{17+0} \times 100\% \quad \text{Then the Recall value} = 100\%$$

Evaluasi metode Naive Bayes pada penelitian ini dilakukan menggunakan Confusion Matrix untuk mengukur tingkat efektivitas model dalam mengklasifikasikan kepuasan pelanggan. Berdasarkan tabel Confusion Matrix, seluruh data yang berjumlah 20 berhasil diklasifikasikan dengan benar, di mana terdapat 17 data kategori Puas yang terklasifikasi dengan tepat dan 3 data kategori Tidak Puas yang juga sesuai dengan kelas sebenarnya. Tidak ditemukan kesalahan klasifikasi, baik berupa False Positive (FP) maupun False Negative (FN), yang menunjukkan bahwa model mampu membedakan kedua kelas dengan sangat baik.

Hasil perhitungan metrik evaluasi menunjukkan performa model yang sangat optimal, dengan nilai akurasi sebesar 100%, presisi 100%, dan recall 100%. Nilai akurasi yang sempurna menandakan bahwa seluruh data uji berhasil diklasifikasikan secara benar, sedangkan nilai presisi dan recall yang maksimal menunjukkan bahwa model tidak menghasilkan kesalahan dalam memprediksi kategori Puas maupun Tidak Puas. Dengan hasil tersebut, dapat disimpulkan bahwa metode Naive Bayes sangat efektif dalam memprediksi tingkat kepuasan pelanggan CS Float berdasarkan data yang digunakan dalam penelitian ini.

3.4. Penerapan Metode K-Nearest Neighbor (KNN)

3.4.1. Pengumpulan Data

Pengumpulan data pada penelitian dengan metode K-Nearest Neighbor (KNN) dilakukan dengan menghimpun data penilaian pelanggan terhadap produk CS Float melalui kuesioner yang disebarakan menggunakan Google Form. Data yang dikumpulkan mencakup atribut waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat, serta kategori kepuasan pelanggan

sebagai label kelas. Data hasil kuesioner tersebut digunakan sebagai dataset utama dalam proses analisis, di mana setiap data pelanggan berperan sebagai representasi objek yang akan dibandingkan tingkat kemiripannya dengan data lain dalam proses klasifikasi menggunakan metode KNN.

Tabel 3. 12. Data Hasil Kuesioner

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Bintang syukuriyah siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Nurul Aida Siregar	Cepat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Irwansyah, S.T	Cepat	Kurang Bagus	Bagus	Murah	Bersih	Puas
Nur indah nasution	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Aini Minta Ito Harahap	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Eka Fitri r	Cepat	Bagus	Kurang Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Mahal	Bersih	Puas
Diana syahfitri	Cepat	Bagus	Bagus	Murah	Kurang Bersih	Puas
dela yuuliyani	Lambat	Kurang Bagus	Bagus	Mahal	Kurang Bersih	Tidak Puas
Amar ma'rub siregar	Lambat	Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Nurhayati	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Bersih	Tidak Puas
putrireimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Grifaril	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Zuhri Dio alfalah	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Semangat abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
IKA MAYA SARI SIMAMORA	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurul aminah fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Hani aura arcika	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas

Dinda cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia rambe	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn syafrida	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn	Cepat	Bagus	Bagus	Murah	Bersih	Puas
dhea allya ginting	Cepat	Bagus	Bagus	Murah	Bersih	Puas
SHOLE HAMID HSB	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas menunjukkan data penelitian yang dikumpulkan dari sebanyak 40 responden melalui penyebaran kuesioner kepada pelanggan produk CS Float. Data tersebut memuat informasi penilaian pelanggan terhadap beberapa atribut, yaitu waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat, serta kategori kepuasan pelanggan sebagai variabel kelas. Seluruh data yang diperoleh mencerminkan persepsi responden terhadap kualitas produk dan layanan yang diberikan, sehingga dapat digunakan sebagai dataset dalam proses analisis dan penerapan metode klasifikasi, khususnya K-Nearest Neighbor (KNN), untuk memprediksi tingkat kepuasan pelanggan secara akurat.

3.4.2. Pembersihan Data

Pembersihan data dilakukan untuk memastikan bahwa data hasil kuesioner yang digunakan dalam penelitian memiliki kualitas yang baik dan layak untuk

dianalisis menggunakan metode K-Nearest Neighbor (KNN). Pada tahap ini, seluruh data responden diperiksa untuk menghilangkan data ganda, memperbaiki kesalahan penulisan, serta memastikan tidak terdapat data yang kosong atau tidak konsisten pada setiap atribut penilaian. Selain itu, penyeragaman format data pada setiap variabel dilakukan agar seluruh data memiliki struktur yang sama dan dapat diproses dengan baik. Data yang telah melalui tahap pembersihan ini selanjutnya digunakan sebagai dataset yang valid dan siap diterapkan dalam proses klasifikasi tingkat kepuasan pelanggan.

Tabel 3. 13. Data Hasil Pembersihan

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Aini Minta Ito Harahap	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Amar Ma'rub Siregar	Lambat	Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Bintang Syukuriyah Siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dela Yuuliyani	Lambat	Kurang Bagus	Bagus	Mahal	Kurang Bersih	Tidak Puas
Dhea Allya Ginting	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Diana Syahfitri	Cepat	Bagus	Bagus	Murah	Kurang Bersih	Puas
Dinda Cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia Rambe	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Eka Fitri R	Cepat	Bagus	Kurang Bagus	Murah	Bersih	Puas
Grifaril	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Hani Aura Arcika	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Ika Maya Sari Simamora	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Irwansyah, S.T	Cepat	Kurang Bagus	Bagus	Murah	Bersih	Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nur Indah Nasution	Lambat	Kurang Bagus	Kurang Bagus	Murah	Kurang Bersih	Tidak Puas
Nurhayati	Lambat	Kurang Bagus	Kurang Bagus	Mahal	Bersih	Tidak Puas
Nurul Aida Siregar	Cepat	Kurang Bagus	Kurang Bagus	Mahal	Kurang Bersih	Tidak Puas
Nurul Aminah Fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Putrireimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Ririn Syafrida	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Mahal	Bersih	Puas
Semangat Abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Shole Hamid Hsb	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria Syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Zuhri Dio Alfalah	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Tabel di atas merupakan data penelitian yang telah melalui proses pembersihan data (data cleaning) sebelum digunakan dalam tahap analisis lebih lanjut. Pada proses ini, seluruh data responden telah diperiksa sehingga tidak terdapat data ganda, kesalahan penulisan, maupun data yang tidak lengkap pada setiap atribut penilaian, seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Selain itu, penulisan nilai pada setiap variabel telah diseragamkan agar konsisten dan mudah diolah. Data yang telah dibersihkan ini selanjutnya digunakan sebagai dataset yang valid dan siap diterapkan dalam

proses klasifikasi tingkat kepuasan pelanggan menggunakan metode K-Nearest Neighbor (KNN).

3.4.3. Transformasi Data

Transformasi data dilakukan untuk mengubah data yang semula berbentuk kategorikal menjadi data numerikal agar dapat diproses oleh metode klasifikasi yang digunakan dalam penelitian ini. Pada tahap ini, setiap kategori pada variabel seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat dikonversi ke dalam bentuk nilai angka sesuai dengan ketentuan yang telah ditetapkan. Proses transformasi ini bertujuan untuk menyederhanakan representasi data, mempermudah perhitungan, serta memungkinkan algoritma seperti Naive Bayes dan K-Nearest Neighbor (KNN) bekerja secara optimal dalam mengenali pola dan melakukan prediksi tingkat kepuasan pelanggan.

Tabel 3. 14. Data Hasil Transformasi

Nama Lengkap	Waktu Respon	Kualiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adelia Putri	10	10	10	10	10	Puas
Adinda Sevira	10	10	10	10	10	Puas
Aini Minta Ito Harahap	5	5	5	10	5	Tidak Puas
Amar Ma'rub Siregar	5	10	5	5	5	Tidak Puas
Arnes Dian Putri Harefa	10	10	10	10	10	Puas
Bintang Syukuriyah Siregar	5	10	10	10	10	Puas
Dea	10	10	10	10	10	Puas
Dela Yuuliyani	5	5	10	5	5	Tidak Puas
Dhea Allya Ginting	10	10	10	10	10	Puas
Diana Syahfitri	10	10	10	10	5	Puas
Dinda Cantiks	10	10	10	10	10	Puas

Dinda Olivia Rambe	5	5	5	5	5	Tidak Puas
Divan Alfaruzi	5	5	5	5	5	Tidak Puas
Eka Fitri R	10	10	5	10	10	Puas
Grifaril	10	10	10	10	10	Puas
Hani Aura Arcika	5	5	5	5	5	Tidak Puas
Husna	10	10	10	10	10	Puas
Ika Maya Sari Simamora	10	10	10	10	10	Puas
Indah Rahmadani	10	10	10	10	10	Puas
Irwansyah, S.T	10	5	10	10	10	Puas
Madzy	10	10	10	10	10	Puas
Nadya Fitriani	10	10	10	10	10	Puas
Nur Indah Nasution	5	5	5	10	5	Tidak Puas
Nurhayati	5	5	5	5	10	Tidak Puas
Nurul Aida Siregar	10	5	5	5	5	Tidak Puas
Nurul Aminah Fajri	10	10	10	10	10	Puas
Putreimadinirasti	10	10	10	10	10	Puas
Ririn	10	10	10	10	10	Puas
Ririn Syafrida	10	10	10	10	10	Puas
Rista Rahayu	10	10	10	10	10	Puas
Rizalunsani Rambe	10	10	10	10	10	Puas
Saima Putri Siregar	10	10	10	5	10	Puas
Semangat Abadi Telaumbanua	10	10	10	10	10	Puas
Shole Hamid Hsb	10	10	10	10	10	Puas
Sindi Rambe	10	10	10	10	10	Puas
Solehuddin Siregar	10	10	10	10	10	Puas
Tika Andini	10	10	10	10	10	Puas
Tria Syahputri	10	10	10	10	10	Puas
Tuti Rama Yani	10	10	10	10	10	Puas
Zuhri Dio Alfalah	10	10	10	10	10	Puas

Tabel di atas merupakan data penelitian yang telah melalui tahap transformasi data, di mana seluruh atribut penilaian yang sebelumnya berbentuk kategorikal telah diubah ke dalam bentuk numerikal. Setiap variabel seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat direpresentasikan dalam bentuk nilai angka untuk memudahkan proses perhitungan

dan analisis menggunakan metode klasifikasi. Transformasi ini bertujuan agar data dapat diolah secara matematis oleh algoritma seperti Naive Bayes dan K-Nearest Neighbor (KNN), sehingga proses pemodelan dan prediksi tingkat kepuasan pelanggan dapat dilakukan secara lebih akurat dan sistematis.

3.4.4. Pembagian Data

Pembagian data pada penelitian ini dilakukan untuk mendukung proses analisis dan pengujian model klasifikasi secara objektif. Data yang telah melalui tahap pembersihan selanjutnya dibagi menjadi dua bagian, yaitu data training dan data testing. Data training digunakan untuk membangun dan melatih model agar mampu mengenali pola hubungan antara atribut penilaian dengan kategori kepuasan pelanggan, sedangkan data testing digunakan untuk menguji kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data ini bertujuan untuk mengevaluasi kinerja model secara akurat serta memastikan hasil prediksi yang dihasilkan dapat dipertanggungjawabkan secara ilmiah.

Tabel 3. 15. Data Training

Nama Lengkap	Waktu Respon	Kualiiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Tuti Rahayu	10	10	10	10	10	Puas

Tabel 3. 16. Data Testing

Nama Lengkap	Waktu Respon	Kualiiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adelia Putri	10	10	10	10	10	Puas
Adinda Sevira	10	10	10	10	10	Puas

Arnes Dian Putri Harefa	10	10	10	10	10	Puas
Dea	10	10	10	10	10	Puas
dhea allya ginting	10	10	10	10	10	Puas
Dinda cantiks	10	10	10	10	10	Puas
Dinda Olivia rambe	5	5	5	5	5	Tidak Puas
Divan Alfaruzi	5	5	5	5	5	Tidak Puas
Hani aura arcika	5	5	5	5	5	Tidak Puas
Husna	10	10	10	10	10	Puas
Indah Rahmadani	10	10	10	10	10	Puas
Madzy	10	10	10	10	10	Puas
Nurul aminah fajri	10	10	10	10	10	Puas
Ririn	10	10	10	10	10	Puas
Ririn syafrida	10	10	10	10	10	Puas
Rizalunsani Rambe	10	10	10	10	10	Puas
SHOLE HAMID HSB	10	10	10	10	10	Puas
Sindi Rambe	10	10	10	10	10	Puas
Solehuddin Siregar	10	10	10	10	10	Puas
Tika Andini	10	10	10	10	10	Puas

Tabel di atas menunjukkan pembagian data penelitian ke dalam data training dan data testing yang digunakan dalam proses analisis. Pada penelitian ini, data training hanya terdiri dari satu data karena disesuaikan dengan kebutuhan pengujian model yang bersifat sederhana dan sebagai acuan awal dalam proses klasifikasi. Data training tersebut berfungsi sebagai representasi pola dasar kategori kepuasan pelanggan, sedangkan data testing digunakan untuk menguji kemampuan model dalam melakukan prediksi terhadap data lainnya. Pembagian data dengan jumlah data training yang terbatas ini dilakukan berdasarkan pertimbangan kebutuhan penelitian, tanpa mengurangi tujuan utama penelitian dalam mengamati hasil klasifikasi tingkat kepuasan pelanggan.

3.4.5. Perhitungan Data

Perhitungan pada metode K-Nearest Neighbor (KNN) dilakukan dengan mencari nilai kedekatan antara data uji dan data latih menggunakan rumus Euclidean Distance pada seluruh atribut numerik yang tersedia. Proses ini dilakukan dengan menghitung selisih tiap atribut, kemudian mengkuadratkannya, menjumlahkan seluruh hasil kuadrat, dan mengambil nilai akarnya untuk memperoleh jarak akhir. Nilai jarak tersebut menunjukkan tingkat kemiripan data, di mana dua data dengan jarak terdekat ($K = 2$) dijadikan dasar penentuan kelas sehingga kategori kepuasan pelanggan dapat diketahui berdasarkan mayoritas tetangga terdekatnya.

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2}$$

Pada penelitian ini untuk perhitungannya penulis akan menggunakan data Adelia Putri. Untuk perhitungannya yaitu sebagai berikut.

$$d(x, y) = \sqrt{(10 - 10)^2 + (10 - 10)^2 + (10 - 10)^2 + (10 - 10)^2 + (10 - 10)^2}$$

$$d(x, y) = \sqrt{(0)^2 + (0)^2 + (0)^2 + (0)^2 + (0)^2}$$

$$d(x, y) = \sqrt{0}$$

$$d(x, y) = 0$$

Untuk hasil perhitungan selanjutnya akan dipaparkan dalam bentuk tabel yaitu sebagai berikut.

Tabel 3. 17. Data Hasil Perhitungan Metode KNN

Nama Lengkap	Perhitungan	Kategori
Adelia Putri	0.00	Puas
Adinda Sevira	0.00	Puas
Arnes Dian Putri Harefa	0.00	Puas
Dea	0.00	Puas

Dhea Allya Ginting	0.00	Puas
Dinda Cantiks	0.00	Puas
Dinda Olivia Rambe	11.18	Tidak Puas
Divan Alfaruzi	11.18	Tidak Puas
Hani Aura Arcika	11.18	Tidak Puas
Husna	0.00	Puas
Indah Rahmadani	0.00	Puas
Madzy	0.00	Puas
Nurul Aminah Fajri	0.00	Puas
Ririn	0.00	Puas
Ririn Syafrida	0.00	Puas
Rizalunsani Rambe	0.00	Puas
Shole Hamid Hsb	0.00	Puas
Sindi Rambe	0.00	Puas
Solehuddin Siregar	0.00	Puas
Tika Andini	0.00	Puas

Tabel di atas merupakan hasil klasifikasi kepuasan pelanggan menggunakan metode K-Nearest Neighbor (KNN) dengan parameter $K = 1$, di mana penentuan kelas dilakukan berdasarkan satu data terdekat. Hasil perhitungan menunjukkan bahwa sebagian besar data memiliki nilai jarak 0,00, yang menandakan bahwa data uji memiliki nilai atribut yang sama persis dengan data latih sehingga tingkat kemiripannya sangat tinggi. Oleh karena itu, seluruh data dengan jarak 0,00 diklasifikasikan ke dalam kategori Puas, sesuai dengan kelas pada data terdekatnya.

Sementara itu, beberapa data seperti Dinda Olivia Rambe, Divan Alfaruzi, dan Hani Aura Arcika memiliki nilai jarak 11,18, yang menunjukkan adanya perbedaan cukup besar antara data uji dan data latih. Karena data terdekatnya berada pada kategori Tidak Puas, maka ketiga data tersebut diklasifikasikan sebagai Tidak Puas. Dengan penggunaan $K = 1$, metode KNN sangat bergantung pada kedekatan satu tetangga terdekat, sehingga hasil klasifikasi secara langsung mencerminkan tingkat kemiripan data berdasarkan jarak Euclidean yang dihitung.

Selain itu, hasil ini juga menunjukkan bahwa penggunaan nilai $K = 1$ membuat model KNN menjadi sangat sensitif terhadap kesamaan data. Ketika data

uji identik dengan data latih, nilai jarak yang dihasilkan adalah 0,00 dan keputusan klasifikasi menjadi sangat pasti. Kondisi ini menguntungkan apabila data yang digunakan konsisten dan representatif, namun di sisi lain berpotensi menimbulkan bias jika terdapat noise atau variasi kecil pada data. Oleh karena itu, meskipun hasil klasifikasi pada penelitian ini terlihat jelas dan tegas, pemilihan nilai K yang lebih besar dapat dipertimbangkan pada penelitian selanjutnya untuk memperoleh hasil klasifikasi yang lebih stabil dan robust.

3.4.6. Evaluasi Model

Evaluasi metode K-Nearest Neighbor (KNN) pada penelitian ini menggunakan Confusion Matrix. Hasil tersebut yang nantinya akan memberikan nilai efektivitas dari metode yang digunakan.

		Predicted		Σ
		Puas	Tidak Puas	
Actual	Puas	17	0	17
	Tidak Puas	0	3	3
Σ		17	3	20

Hasil True Positive (TP) adalah 17. True Negative (TN) adalah 3, False Positive (FP) adalah 0 dan False Negative (FN) adalah 0. Maka Nilai akurasi, presisi dan recall adalah sebagai berikut:

$$\text{Accuracy} = \frac{17+3}{17+3+0+0} \times 100\% \quad \text{Then the Accuracy value} = 100\%$$

$$\text{Presisi} = \frac{17}{17+0} \times 100\% \quad \text{Then the Precision value} = 100\%$$

$$\text{Recall} = \frac{17}{17+0} \times 100\% \quad \text{Then the Recall value} = 100\%$$

Evaluasi metode K-Nearest Neighbor (KNN) pada penelitian ini dilakukan menggunakan Confusion Matrix untuk menilai tingkat keefektifan model dalam mengklasifikasikan kepuasan pelanggan. Berdasarkan hasil Confusion Matrix, seluruh data uji yang berjumlah 20 data berhasil diklasifikasikan dengan tepat, di mana 17 data dengan kategori Puas dan 3 data dengan kategori Tidak Puas seluruhnya terprediksi sesuai dengan kelas sebenarnya. Tidak ditemukan kesalahan klasifikasi, baik False Positive (FP) maupun False Negative (FN), yang menunjukkan bahwa model KNN mampu mengelompokkan data dengan sangat akurat.

Perhitungan metrik evaluasi menunjukkan bahwa metode KNN memperoleh nilai akurasi sebesar 100%, presisi 100%, dan recall 100%. Nilai akurasi yang sempurna menandakan bahwa seluruh data uji berhasil diprediksi dengan benar, sementara nilai presisi dan recall yang maksimal menunjukkan kemampuan model dalam mengidentifikasi kategori kepuasan tanpa menghasilkan kesalahan prediksi. Dengan demikian, dapat disimpulkan bahwa metode KNN dengan parameter yang digunakan dalam penelitian ini sangat efektif dalam memprediksi tingkat kepuasan pelanggan CS Float.

3.5. Perbandingan Hasil dari Dua Model Klasifikasi

Berdasarkan hasil evaluasi menggunakan Confusion Matrix, metode Naive Bayes dan K-Nearest Neighbor (KNN) menunjukkan performa yang sama-sama sangat optimal dalam mengklasifikasikan tingkat kepuasan pelanggan CS Float. Kedua metode berhasil mengklasifikasikan seluruh data uji yang berjumlah 20 data secara tepat tanpa adanya kesalahan klasifikasi. Hal ini ditunjukkan oleh tidak

ditemukannya nilai False Positive (FP) maupun False Negative (FN) pada kedua model, sehingga seluruh data kategori Puas dan Tidak Puas dapat diprediksi sesuai dengan kelas sebenarnya.

Dari sisi metrik evaluasi, baik Naive Bayes maupun KNN menghasilkan nilai akurasi, presisi, dan recall sebesar 100%. Kesamaan nilai ini menunjukkan bahwa kedua metode memiliki kemampuan yang setara dalam mengenali pola kepuasan pelanggan berdasarkan data yang digunakan. Naive Bayes menunjukkan keunggulan dalam pendekatan probabilistik yang sederhana namun efektif, sedangkan KNN mengandalkan kedekatan jarak antar data untuk menentukan kelas, namun keduanya mampu menghasilkan tingkat ketepatan klasifikasi yang maksimal pada dataset penelitian ini.

Meskipun hasil evaluasi menunjukkan performa yang identik, perbedaan mendasar antara kedua metode tetap terlihat pada mekanisme kerjanya. Naive Bayes cenderung lebih efisien dalam proses komputasi karena menggunakan perhitungan probabilitas, sementara KNN membutuhkan perhitungan jarak terhadap data latih sehingga dapat menjadi lebih kompleks ketika jumlah data meningkat. Dengan demikian, kedua metode terbukti sama-sama efektif dalam penelitian ini, namun pemilihan metode yang paling sesuai dapat disesuaikan dengan karakteristik data dan kebutuhan sistem prediksi kepuasan pelanggan CS Float di masa mendatang.