Journal of Computer Science and Information Systems (JCoInS)
Program Studi Sistem Informasi, Fakultas Sains & Teknologi, Universitas Labuhanbatu
Volume : 7, Nomor : 3, Juli 2026, Pages. 175-191
e-ISSN: 2747-2221

Penerapan Algoritma *Naïve Bayes Classifier* dan *Decision Tree* untuk Memprediksi Tingkat Kepuasan Pelanggan Teras Coffe Rantauprapat

[1]Elfi Zahra Yuni, [2]Syaiful Zuhri Harahap, [3]Ibnu Rasyid Munthe, [4]Angga Putra Juledi

[1,2,3,4]Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

Email: [1]elfizahrayunii@gmail.com, [2]syaifulzuhriharahap@gmail.com, [3]ibnurasyidmunthe@gmail.com, [4]anggapj19@gmail.com

Corresponding Author : elfizahrayunii@gmail.com

Abstract

The increasingly fierce competition in the coffee shop business requires business owners to understand customer satisfaction levels as a basis for improving service quality and maintaining customer loyalty. Therefore, an approach capable of accurately identifying and predicting customer satisfaction levels based on customer data is needed. Data mining is a data processing technique that can be used to discover patterns and important information from data sets to support the decision-making process. In this study, the Naïve Bayes Classifier and Decision Tree algorithms were used because both are classification methods capable of generating predictions based on the characteristics of the data. The research method used was a quantitative method by utilizing Teras Coffee Rantauprapat customer questionnaire data which was then processed using the Orange Data Mining application. The research data was divided into training data and testing data to build and test the classification models generated by both algorithms. The results showed that the Naïve Bayes and Decision Tree algorithms were able to classify customer satisfaction levels into satisfied and dissatisfied categories with a good level of accuracy. Based on the model evaluation results, the Naïve Bayes algorithm obtained superior performance compared to Decision Tree based on higher AUC, Precision, F1-Score, and MCC values. Thus, both algorithms can be applied to predict customer satisfaction levels, but Naïve Bayes proved more optimal in generating predictions on the dataset used in this study. The results of this study are expected to serve as a reference for Teras Coffee Rantauprapat in continuously improving service quality and customer satisfaction.

Keywords: *Data Mining, Decision Trees, Customer Satisfaction, Naïve Bayes Classifier, Orange Data Mining.*

1. Pendahuluan

Peningkatan persaingan bisnis pada sektor kuliner dan kedai kopi di Indonesia mendorong setiap pelaku usaha untuk terus meningkatkan kualitas pelayanan dan kepuasan pelanggan. Kepuasan pelanggan merupakan salah satu indikator penting yang menentukan keberlangsungan usaha karena pelanggan yang merasa puas cenderung melakukan pembelian ulang serta merekomendasikan produk atau layanan kepada orang lain. Di tengah berkembangnya tren konsumsi kopi yang semakin meningkat, berbagai kedai kopi berlomba-lomba menawarkan produk berkualitas, pelayanan yang baik, serta suasana yang nyaman guna menarik minat konsumen. Kondisi tersebut menyebabkan pelaku usaha tidak hanya dituntut untuk mampu menyediakan produk yang sesuai dengan

kebutuhan pelanggan, tetapi juga harus memahami tingkat kepuasan pelanggan secara lebih mendalam. Oleh karena itu, pengukuran dan prediksi kepuasan pelanggan menjadi aspek yang penting untuk dilakukan sebagai dasar dalam pengambilan keputusan bisnis yang lebih efektif. Teras Coffee Rantauprapat merupakan salah satu usaha yang bergerak di bidang penyediaan makanan dan minuman yang menghadapi tantangan dalam mempertahankan loyalitas pelanggan di tengah persaingan yang semakin kompetitif. Keberhasilan usaha tidak hanya ditentukan oleh kualitas produk yang ditawarkan, tetapi juga dipengaruhi oleh berbagai faktor lain seperti kualitas pelayanan, kenyamanan tempat, kebersihan lingkungan, kecepatan pelayanan, serta kesesuaian harga dengan harapan pelanggan. Beragam faktor tersebut dapat memengaruhi persepsi pelanggan terhadap layanan yang diterima sehingga menghasilkan tingkat kepuasan yang berbeda-beda. Selama ini, evaluasi kepuasan pelanggan sering kali dilakukan secara sederhana melalui pengamatan langsung atau masukan yang bersifat subjektif. Pendekatan tersebut dinilai belum mampu memberikan gambaran yang komprehensif mengenai pola kepuasan pelanggan secara keseluruhan sehingga diperlukan metode yang lebih sistematis dalam menganalisis data pelanggan. Perkembangan teknologi informasi dan ilmu pengetahuan telah mendorong pemanfaatan data dalam berbagai bidang, termasuk dalam analisis perilaku konsumen. Salah satu pendekatan yang banyak digunakan untuk menggali informasi dari kumpulan data adalah data mining. Data mining memungkinkan proses identifikasi pola, hubungan, serta pengetahuan baru yang sebelumnya tersembunyi dalam suatu basis data. Dalam penelitian ini, proses pengolahan data dilakukan menggunakan pendekatan *Knowledge Discovery in Database* (KDD) yang terdiri atas beberapa tahapan, yaitu seleksi data, pembersihan data, transformasi data, proses data mining, dan evaluasi hasil. Melalui tahapan tersebut, data hasil kuesioner pelanggan dapat diolah menjadi informasi yang bermanfaat untuk mengetahui faktor-faktor yang memengaruhi tingkat kepuasan pelanggan. Dengan demikian, pemanfaatan data mining diharapkan mampu menghasilkan informasi yang lebih objektif dan akurat dibandingkan metode analisis konvensional. Salah satu teknik yang sering digunakan dalam data mining adalah metode klasifikasi, yaitu proses pengelompokan data berdasarkan kategori tertentu. Dalam penelitian ini digunakan dua algoritma klasifikasi, yaitu Naïve Bayes Classifier dan Decision Tree (Pratama, Yanris, Nirmala, & Hasibuan, 2023) (Ayuningtyas, Nining, & Basysyar, 2022) (Melyani & Harahap, 2024) (Aditya, Munthe, & Sihombing, 2024). Naïve Bayes Classifier dikenal sebagai algoritma yang memiliki kemampuan klasifikasi yang baik dengan proses perhitungan yang relatif sederhana serta mampu bekerja secara efektif pada berbagai jenis data. Sementara itu, Decision Tree memiliki keunggulan dalam menghasilkan model yang mudah dipahami karena disajikan dalam bentuk struktur pohon keputusan yang menggambarkan hubungan antarvariabel secara jelas (Wijaya, Dharma, Heyneker, & Vanness, 2023). Penggunaan kedua algoritma tersebut memungkinkan peneliti untuk membandingkan performa masing-masing metode dalam memprediksi tingkat kepuasan pelanggan berdasarkan data yang diperoleh. Hasil perbandingan tersebut dapat memberikan gambaran mengenai algoritma yang memiliki tingkat akurasi lebih baik dalam mendukung pengambilan keputusan. Berdasarkan uraian tersebut, penelitian mengenai penerapan algoritma Naïve Bayes Classifier dan Decision Tree untuk memprediksi tingkat kepuasan pelanggan pada Teras Coffee Rantauprapat menjadi penting untuk dilakukan (Nugraha & Romadhony, 2023) (Juliadi & Puspitarani, 2022) (Zai, Sirait, Nainggolan, Sihombing, & Banjarnahor, 2023) (Ramadhon, Ogi, & Agung, 2024). Penelitian ini diharapkan dapat menghasilkan model prediksi yang mampu mengidentifikasi tingkat kepuasan pelanggan secara lebih akurat berdasarkan data yang

diperoleh dari hasil survei. Selain itu, penelitian ini juga diharapkan dapat memberikan informasi mengenai algoritma yang memiliki performa terbaik dalam proses klasifikasi data kepuasan pelanggan. Hasil penelitian tidak hanya bermanfaat bagi pengembangan ilmu pengetahuan di bidang data mining dan machine learning, tetapi juga dapat menjadi bahan pertimbangan bagi pihak Teras Coffee Rantauprapat dalam merumuskan strategi peningkatan kualitas pelayanan dan kepuasan pelanggan. Dengan demikian, penelitian ini diharapkan mampu memberikan kontribusi nyata dalam mendukung keberhasilan dan keberlanjutan usaha di masa yang akan datang.

2. Landasan Teori

Data Mining

Data mining merupakan suatu proses pengolahan dan analisis data yang bertujuan untuk menemukan pola, hubungan, tren, serta informasi tersembunyi dari sejumlah data yang berukuran besar sehingga dapat menghasilkan pengetahuan yang bermanfaat bagi pengambilan keputusan (Diansyah, 2022). Data mining menjadi salah satu bagian penting dalam perkembangan teknologi informasi karena mampu mengubah data mentah menjadi informasi yang memiliki nilai guna bagi organisasi maupun perusahaan. Teknik ini menggabungkan berbagai disiplin ilmu seperti statistika, kecerdasan buatan (*artificial intelligence*), pembelajaran mesin (*machine learning*), dan basis data untuk menghasilkan model analisis yang akurat (Diana Dewi, Qisthi, Lestari, & Putri, 2023). Dalam penerapannya, data mining banyak digunakan untuk kegiatan klasifikasi, prediksi, klasterisasi, dan pencarian pola asosiasi pada berbagai bidang, termasuk bisnis, pendidikan, kesehatan, dan pemasaran (Anwar, Ambiyar, & Fadhilah, 2023). Oleh karena itu, data mining menjadi pendekatan yang efektif dalam membantu perusahaan memahami karakteristik pelanggan, memprediksi perilaku konsumen, serta mendukung penyusunan strategi yang lebih tepat berdasarkan hasil analisis data yang tersedia.

Metode Naïve Bayes

Naïve Bayes Classifier merupakan salah satu algoritma klasifikasi yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap atribut atau variabel yang digunakan bersifat independen terhadap atribut lainnya (Apriyani, Maskuri, Ratsanjani, Pramudhita, & Rawansyah, 2023) (Adjani, Fauzia, & Juliane, 2023). Algoritma ini bekerja dengan menghitung probabilitas kemunculan suatu kelas berdasarkan data yang telah diketahui sebelumnya, kemudian menentukan kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi. Naïve Bayes dikenal sebagai metode yang sederhana, cepat, dan efisien dalam mengolah data berukuran besar, sehingga banyak digunakan dalam berbagai bidang seperti klasifikasi dokumen, deteksi spam, analisis sentimen, hingga prediksi kepuasan pelanggan (Siregar, Irmayani, & Sari, 2023) (Wahyudi et al., 2022). Meskipun menggunakan asumsi independensi yang dalam kondisi nyata sering kali tidak sepenuhnya terpenuhi, algoritma ini tetap mampu menghasilkan tingkat akurasi yang baik pada berbagai kasus klasifikasi. Oleh karena itu, Naïve Bayes menjadi salah satu metode yang banyak dipilih dalam penelitian data mining dan machine learning karena kemudahan implementasi serta kemampuannya dalam menghasilkan prediksi yang efektif (Hasri & Alita, 2022).

Metode *Decision Tree*

Decision Tree merupakan salah satu metode klasifikasi dalam data mining dan machine learning yang bekerja dengan membentuk struktur pohon keputusan untuk

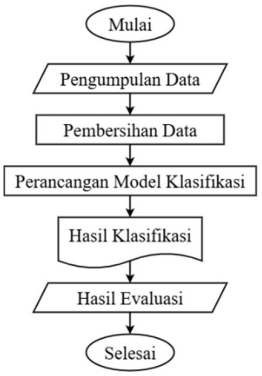
memprediksi suatu kelas atau kategori berdasarkan atribut-atribut yang dimiliki data (Nugraha & Romadhony, 2023). Metode ini membagi data ke dalam beberapa cabang berdasarkan nilai atribut yang dianggap paling berpengaruh, sehingga menghasilkan aturan keputusan yang mudah dipahami dan diinterpretasikan. Setiap simpul (*node*) pada pohon keputusan merepresentasikan atribut pengujian, sedangkan cabang menunjukkan hasil dari pengujian tersebut dan daun (*leaf*) menunjukkan hasil klasifikasi akhir (Anam, Nurhakim, & Juliane, 2022).

3. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif dengan pendekatan klasifikasi data untuk memprediksi tingkat kepuasan pelanggan pada Teras Coffee Rantauprapat. Metode kuantitatif dipilih karena penelitian ini menggunakan data yang diperoleh melalui penyebaran kuesioner kepada pelanggan, kemudian diolah dan dianalisis secara statistik untuk menghasilkan informasi yang objektif dan terukur.

Dalam proses analisis data, penelitian ini menggunakan dua metode klasifikasi, yaitu Naïve Bayes Classifier dan Decision Tree (Alam, Alana, & Juliane, 2023). Naïve Bayes Classifier merupakan metode yang bekerja berdasarkan konsep probabilitas untuk menentukan kemungkinan suatu data termasuk ke dalam kategori tertentu berdasarkan karakteristik yang dimilikinya. Sementara itu, Decision Tree merupakan metode yang membentuk struktur pohon keputusan untuk mengklasifikasikan data berdasarkan atribut-atribut yang paling berpengaruh terhadap hasil prediksi. Penggunaan kedua metode tersebut bertujuan untuk menghasilkan model prediksi tingkat kepuasan pelanggan serta membandingkan performa masing-masing algoritma dalam mengolah data yang diperoleh. Melalui perbandingan tersebut, dapat diketahui metode yang memiliki tingkat akurasi lebih baik sehingga mampu memberikan hasil prediksi yang lebih optimal dalam mendukung evaluasi dan peningkatan kualitas pelayanan pada Teras Coffee Rantauprapat.

Kerangka Kerja Penelitian



Gambar 1. Kerangka Kerja Penelitian

4. Hasil Dan Pembahasan

Pengumpulan Data

Pengumpulan data merupakan tahap awal dalam penelitian yang dilakukan untuk memperoleh informasi yang dibutuhkan sebagai dasar analisis dan pengolahan data. Pada penelitian ini, data dikumpulkan melalui penyebaran kuesioner kepada pelanggan Teras Coffee Rantauprapat guna memperoleh informasi mengenai tingkat kepuasan pelanggan

berdasarkan berbagai aspek yang telah ditentukan. Data yang diperoleh dari responden kemudian digunakan sebagai sumber utama dalam proses klasifikasi menggunakan algoritma Naïve Bayes Classifier dan Decision Tree. Tahap pengumpulan data dilakukan secara sistematis untuk memastikan data yang diperoleh sesuai dengan tujuan penelitian dan mampu merepresentasikan kondisi sebenarnya. Oleh karena itu, kualitas data yang dikumpulkan sangat berpengaruh terhadap ketepatan hasil analisis dan prediksi yang dihasilkan dalam penelitian.

Tabel 1. Data Sampel

Nama	Kualitas Pelayanan	Kualitas Produk	Kenyamanan Tempat	Kebersihan
Agus Setiawan	Baik	Tidak Baik	Baik	Baik
Aldi Akbar	Baik	Baik	Tidak Baik	Baik
Aldi Kartika	Baik	Baik	Tidak Baik	Baik
Aldi Kurniawan	Baik	Baik	Baik	Baik
Aldi Pertiwi	Baik	Baik	Tidak Baik	Baik
Andi Aditya	Baik	Tidak Baik	Baik	Baik
Andi Hartono	Baik	Tidak Baik	Baik	Baik
Andi Ramadhan	Baik	Baik	Tidak Baik	Baik
Anisa Oktaviani	Baik	Baik	Tidak Baik	Tidak Baik
Anisa Pertiwi	Tidak Baik	Baik	Tidak Baik	Baik
Arif Fikri	Baik	Baik	Tidak Baik	Baik
Arif Handayani	Baik	Baik	Baik	Baik
Ayu Hartono	Baik	Baik	Baik	Baik
Bagas Fauzi	Tidak Baik	Baik	Baik	Baik
Bagas Setiawan	Baik	Tidak Baik	Tidak Baik	Tidak Baik
Bayu Pertiwi	Baik	Baik	Baik	Baik
Budi Puspita	Tidak Baik	Baik	Baik	Tidak Baik
Citra Akbar	Tidak Baik	Tidak Baik	Baik	Baik
Citra Dewi	Baik	Baik	Tidak Baik	Tidak Baik
Citra Fikri	Baik	Baik	Baik	Baik
Citra Lestari	Baik	Baik	Tidak Baik	Baik
Denny Hidayat	Baik	Baik	Tidak Baik	Baik
Dimas Akbar	Baik	Baik	Tidak Baik	Baik
Dimas Santoso	Baik	Baik	Baik	Tidak Baik
Dinda Dewi	Baik	Baik	Baik	Baik
Dinda Fadilah	Baik	Baik	Baik	Baik
Dinda Setiawan	Baik	Baik	Baik	Tidak Baik
Doni Hidayat	Tidak Baik	Tidak Baik	Baik	Baik
Fajar Irawan	Tidak Baik	Baik	Baik	Baik
Fajar Puspita	Baik	Baik	Tidak Baik	Tidak Baik
Fajar Zahra	Baik	Baik	Tidak Baik	Baik
Farhan Gunawan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Baik
Fikri Gunawan	Baik	Baik	Baik	Baik
Fikri Nugroho	Baik	Baik	Baik	Tidak Baik
Fina Gunawan	Baik	Baik	Tidak Baik	Tidak Baik
Fina Maharani	Baik	Baik	Baik	Baik
Hendra Dewi	Baik	Baik	Baik	Baik
Intan Syah	Baik	Tidak Baik	Baik	Baik
Karina Nugroho	Baik	Baik	Baik	Baik
Kevin Kurniawan	Baik	Baik	Baik	Baik
Kevin Nasution	Baik	Tidak Baik	Baik	Baik
Kevin Permata	Baik	Tidak Baik	Baik	Baik

Kevin Puspita	Baik	Baik	Baik	Baik
Lala Gunawan	Baik	Tidak Baik	Baik	Baik
Lala Hartono	Baik	Baik	Baik	Baik
Lala Permata	Tidak Baik	Baik	Baik	Baik
Lilis Hartono	Baik	Baik	Baik	Baik
Maya Maulana	Baik	Baik	Baik	Baik
Maya Rambe	Baik	Baik	Tidak Baik	Baik
Melati Kurniawan	Baik	Tidak Baik	Baik	Baik
Melati Santoso	Baik	Baik	Baik	Baik
Mutiara Handayani	Baik	Baik	Baik	Baik
Nadia Maulana	Tidak Baik	Baik	Baik	Baik
Nadia Pertiwi	Tidak Baik	Tidak Baik	Baik	Baik
Nadia Rahmah	Baik	Tidak Baik	Baik	Tidak Baik
Nanda Kartika	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Baik
Nanda Oktaviani	Baik	Baik	Tidak Baik	Baik
Novi Azzahra	Tidak Baik	Baik	Tidak Baik	Baik
Novi Puspita	Baik	Baik	Baik	Baik
Novi Putri	Baik	Baik	Baik	Baik
Novi Rambe	Tidak Baik	Baik	Baik	Baik
Putri Handayani	Baik	Tidak Baik	Tidak Baik	Baik
Putri Santoso	Baik	Baik	Tidak Baik	Baik
Rafi Nuraini	Baik	Baik	Baik	Baik
Rahma Anggraini	Baik	Baik	Baik	Tidak Baik
Rahma Irawan	Baik	Baik	Baik	Baik
Rahma Lestari	Baik	Baik	Baik	Baik
Rahma Lubis	Baik	Baik	Tidak Baik	Baik
Rani Kurniawan	Tidak Baik	Baik	Baik	Baik
Rani Wijaya	Baik	Tidak Baik	Baik	Baik
Rayhan Setiawan	Baik	Baik	Tidak Baik	Baik
Reza Fauzi	Baik	Baik	Baik	Baik
Reza Lubis	Baik	Baik	Baik	Baik
Riko Pratama	Tidak Baik	Baik	Baik	Tidak Baik
Rina Nasution	Baik	Tidak Baik	Baik	Tidak Baik
Rio Dewi	Baik	Tidak Baik	Baik	Tidak Baik
Rio Firmansyah	Baik	Tidak Baik	Tidak Baik	Tidak Baik
Rio Lubis	Tidak Baik	Tidak Baik	Baik	Baik
Rizky Lubis	Baik	Baik	Baik	Baik
Salsa Maulana	Baik	Baik	Baik	Tidak Baik
Sari Fauzi	Baik	Baik	Baik	Baik
Sari Saputro	Tidak Baik	Tidak Baik	Baik	Baik
Septian Hartono	Baik	Baik	Baik	Tidak Baik
Sinta Pertiwi	Baik	Baik	Baik	Baik
Surya Handayani	Tidak Baik	Tidak Baik	Baik	Baik
Surya Salsabila	Baik	Tidak Baik	Tidak Baik	Baik
Surya Saputra	Tidak Baik	Baik	Baik	Baik
Surya Wijaya	Baik	Tidak Baik	Baik	Tidak Baik
Vina Akbar	Tidak Baik	Tidak Baik	Baik	Baik
Vina Lestari	Tidak Baik	Baik	Baik	Baik
Vina Maharani	Baik	Baik	Baik	Tidak Baik
Vina Syah	Baik	Baik	Baik	Baik

Wahyu Oktaviani	Baik	Baik	Baik	Tidak Baik
Yoga Zahra	Tidak Baik	Baik	Baik	Baik
Yudi Rambe	Baik	Baik	Baik	Tidak Baik
Yudi Sari	Tidak Baik	Baik	Baik	Baik
Yudi Syah	Tidak Baik	Tidak Baik	Baik	Tidak Baik
Yuliana Irawan	Tidak Baik	Baik	Baik	Baik
Yuliana Oktarina	Baik	Tidak Baik	Tidak Baik	Baik
Yuni Oktarina	Baik	Tidak Baik	Baik	Baik

Pada tabel di atas disajikan data sampel penelitian yang terdiri dari 100 responden pelanggan Teras Coffee Rantauprapat. Data tersebut merupakan hasil pengumpulan informasi yang digunakan sebagai dasar dalam proses analisis menggunakan metode Machine Learning. Setiap responden memberikan penilaian terhadap empat variabel, yaitu kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan. Masing-masing variabel memiliki dua kategori penilaian, yaitu Baik dan Tidak Baik. Penyusunan data dilakukan secara sistematis agar memudahkan proses pengolahan pada tahapan berikutnya. Variasi penilaian yang diberikan oleh setiap responden menunjukkan adanya perbedaan persepsi terhadap kualitas layanan yang diterima. Keberagaman kombinasi nilai pada setiap atribut menjadi informasi penting dalam menemukan pola hubungan antarvariabel. Jumlah sampel sebanyak 100 data dinilai cukup untuk memberikan gambaran mengenai kondisi penilaian pelanggan terhadap layanan yang diberikan. Data tersebut juga telah disusun dalam format yang terstruktur sehingga mudah diolah menggunakan perangkat lunak analisis. Dengan demikian, tabel di atas menjadi sumber data utama yang digunakan dalam keseluruhan proses penelitian. Data yang telah dikumpulkan selanjutnya digunakan sebagai masukan dalam proses pengolahan menggunakan algoritma Naïve Bayes Classifier dan Decision Tree. Keempat variabel yang digunakan berperan sebagai atribut prediktor untuk membentuk model klasifikasi tingkat kepuasan pelanggan. Sebelum dilakukan proses klasifikasi, data akan melalui beberapa tahapan, seperti pembagian data menjadi data training dan data testing, transformasi data, serta proses perhitungan sesuai dengan algoritma yang digunakan. Banyaknya data yang tersedia memberikan kesempatan bagi model untuk mempelajari pola hubungan antarvariabel secara lebih baik. Variasi kombinasi nilai pada setiap atribut juga membantu algoritma dalam mengenali karakteristik data secara lebih akurat. Hasil pengolahan data nantinya akan menunjukkan pola klasifikasi berdasarkan karakteristik penilaian yang diberikan oleh pelanggan. Model yang dihasilkan diharapkan mampu memprediksi tingkat kepuasan pelanggan dengan tingkat akurasi yang baik. Informasi tersebut dapat dimanfaatkan sebagai bahan evaluasi dalam meningkatkan kualitas pelayanan, produk, kenyamanan tempat, dan kebersihan. Semakin baik kualitas data yang digunakan, maka semakin baik pula hasil klasifikasi yang diperoleh. Oleh karena itu, data sampel ini memiliki peranan yang sangat penting dalam mendukung keberhasilan proses analisis dan pencapaian tujuan penelitian.

Pembagian Data

Pembagian data merupakan tahap yang dilakukan untuk memisahkan dataset menjadi data pelatihan (*training data*) dan data pengujian (*testing data*) sebelum proses klasifikasi dilakukan. Data pelatihan digunakan untuk membangun dan melatih model klasifikasi agar dapat mengenali pola hubungan antara atribut yang digunakan dengan tingkat kepuasan pelanggan, sedangkan data pengujian digunakan untuk mengukur

kemampuan model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data dilakukan untuk memastikan bahwa model yang dihasilkan tidak hanya mampu bekerja dengan baik pada data pelatihan, tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru. Dengan demikian, hasil evaluasi yang diperoleh dapat memberikan gambaran yang lebih objektif mengenai kinerja algoritma Naïve Bayes Classifier dan Decision Tree dalam memprediksi tingkat kepuasan pelanggan.

Tabel 1. Data Training

Nama	Kualitas Pelayanan	Kualitas Produk	Kenyamanan Tempat	Kebersihan	Kepuasan Pelanggan
Anisa Pertiwi	Tidak Baik	Baik	Tidak Baik	Baik	Tidak Puas
Rani Wijaya	Baik	Tidak Baik	Baik	Baik	Puas
Vina Lestari	Tidak Baik	Baik	Baik	Baik	Puas
Anisa Oktaviani	Baik	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Andi Ramadhan	Baik	Baik	Tidak Baik	Baik	Puas
Kevin Puspita	Baik	Baik	Baik	Baik	Puas
Putri Handayani	Baik	Tidak Baik	Tidak Baik	Baik	Tidak Puas
Surya Wijaya	Baik	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Rizky Lubis	Baik	Baik	Baik	Baik	Puas
Mutiara Handayani	Baik	Baik	Baik	Baik	Puas
Kevin Permata	Baik	Tidak Baik	Baik	Baik	Puas
Fajar Irawan	Tidak Baik	Baik	Baik	Baik	Puas
Nadia Pertiwi	Tidak Baik	Tidak Baik	Baik	Baik	Tidak Puas
Rio Lubis	Tidak Baik	Tidak Baik	Baik	Baik	Tidak Puas
Melati Santoso	Baik	Baik	Baik	Baik	Puas
Vina Maharani	Baik	Baik	Baik	Tidak Baik	Puas
Dinda Dewi	Baik	Baik	Baik	Baik	Puas
Dimas Akbar	Baik	Baik	Tidak Baik	Baik	Puas
Surya Salsabila	Baik	Tidak Baik	Tidak Baik	Baik	Tidak Puas
Denny Hidayat	Baik	Baik	Tidak Baik	Baik	Puas
Lala Permata	Tidak Baik	Baik	Baik	Baik	Puas
Melati Kurniawan	Baik	Tidak Baik	Baik	Baik	Puas
Sinta Pertiwi	Baik	Baik	Baik	Baik	Puas
Intan Syah	Baik	Tidak Baik	Baik	Baik	Puas
Nanda Kartika	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Novi Puspita	Baik	Baik	Baik	Baik	Puas
Aldi Akbar	Baik	Baik	Tidak Baik	Baik	Puas
Nadia Rahmah	Baik	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Vina Akbar	Tidak Baik	Tidak Baik	Baik	Baik	Tidak Puas
Aldi Pertiwi	Baik	Baik	Tidak Baik	Baik	Puas

Pada tabel di atas disajikan data training penelitian yang berjumlah 30 data responden pelanggan Teras Coffee Rantauprapat. Data training merupakan sekumpulan data yang digunakan sebagai dasar pembelajaran bagi algoritma Decision Tree untuk membentuk model klasifikasi. Setiap data terdiri dari empat variabel prediktor, yaitu kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan, serta satu variabel target berupa kepuasan pelanggan. Masing-masing variabel prediktor memiliki dua kategori penilaian, yaitu Baik dan Tidak Baik, sedangkan variabel target terdiri atas kategori Puas dan Tidak Puas. Penyusunan data dilakukan secara sistematis agar memudahkan proses pengolahan dan analisis. Variasi kombinasi nilai pada setiap atribut menunjukkan bahwa setiap responden memiliki karakteristik penilaian yang berbeda-beda. Keberagaman data tersebut sangat penting karena memungkinkan algoritma

mempelajari hubungan antara setiap atribut dengan tingkat kepuasan pelanggan. Jumlah data training sebanyak 30 responden dinilai cukup untuk membentuk pola klasifikasi awal yang akan digunakan pada proses pembentukan pohon keputusan. Data pada tabel di atas menjadi dasar dalam menghitung nilai entropy dan information gain pada setiap atribut. Oleh karena itu, kualitas data training sangat menentukan hasil klasifikasi yang akan dihasilkan oleh metode Decision Tree. Data training tersebut selanjutnya digunakan untuk menentukan atribut yang memiliki kemampuan terbaik dalam membedakan kategori Puas dan Tidak Puas. Proses pembelajaran dilakukan dengan menghitung distribusi setiap atribut terhadap variabel target sehingga diperoleh nilai entropy dan information gain. Hasil perhitungan tersebut akan digunakan untuk menentukan node akar dan node berikutnya dalam pembentukan pohon keputusan. Semakin baik variasi data yang dimiliki, maka semakin mudah algoritma mengenali pola hubungan antarvariabel. Data training juga berfungsi sebagai acuan dalam menyusun aturan klasifikasi yang akan diterapkan pada data baru. Pola yang terbentuk dari data ini menjadi dasar bagi algoritma dalam melakukan proses prediksi secara otomatis. Keberadaan data dengan kategori Puas maupun Tidak Puas memberikan keseimbangan informasi sehingga model dapat mempelajari kedua kelas dengan lebih baik. Selain itu, variasi nilai pada setiap atribut membantu menghasilkan aturan keputusan yang lebih representatif terhadap kondisi sebenarnya. Model yang dibangun dari data training diharapkan mampu memberikan hasil klasifikasi yang akurat ketika diuji menggunakan data testing. Dengan demikian, data training pada tabel di atas menjadi tahapan penting dalam membangun model Decision Tree yang mampu memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat secara efektif.

Tabel 2. Data Testing

Nama	Kualitas Pelayanan	Kualitas Produk	Kenyamanan Tempat	Kebersihan
Fajar Zahra	Baik	Baik	Tidak Baik	Baik
Fikri Gunawan	Baik	Baik	Baik	Baik
Yoga Zahra	Tidak Baik	Baik	Baik	Baik
Kevin Nasution	Baik	Tidak Baik	Baik	Baik
Septian Hartono	Baik	Baik	Baik	Tidak Baik
Surya Saputra	Tidak Baik	Baik	Baik	Baik
Agus Setiawan	Baik	Tidak Baik	Baik	Baik
Kevin Kurniawan	Baik	Baik	Baik	Baik
Rahma Irawan	Baik	Baik	Baik	Baik
Nanda Oktaviani	Baik	Baik	Tidak Baik	Baik
Fina Gunawan	Baik	Baik	Tidak Baik	Tidak Baik
Rio Firmansyah	Baik	Tidak Baik	Tidak Baik	Tidak Baik
Aldi Kurniawan	Baik	Baik	Baik	Baik
Citra Lestari	Baik	Baik	Tidak Baik	Baik
Rio Dewi	Baik	Tidak Baik	Baik	Tidak Baik
Dinda Setiawan	Baik	Baik	Baik	Tidak Baik
Wahyu Oktaviani	Baik	Baik	Baik	Tidak Baik
Novi Azzahra	Tidak Baik	Baik	Tidak Baik	Baik
Andi Aditya	Baik	Tidak Baik	Baik	Baik
Budi Puspita	Tidak Baik	Baik	Baik	Tidak Baik
Lala Hartono	Baik	Baik	Baik	Baik
Aldi Kartika	Baik	Baik	Tidak Baik	Baik
Lilis Hartono	Baik	Baik	Baik	Baik

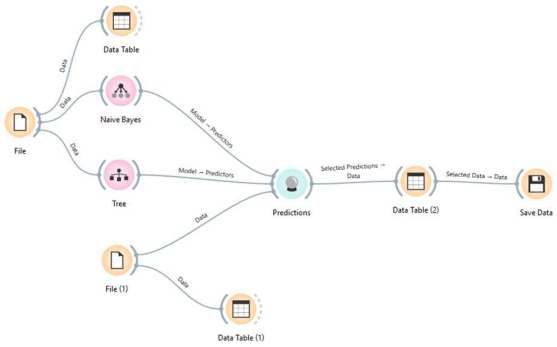
Yudi Rambe	Baik	Baik	Baik	Tidak Baik
Reza Lubis	Baik	Baik	Baik	Baik
Maya Maulana	Baik	Baik	Baik	Baik
Bagas Setiawan	Baik	Tidak Baik	Tidak Baik	Tidak Baik
Dimas Santoso	Baik	Baik	Baik	Tidak Baik
Doni Hidayat	Tidak Baik	Tidak Baik	Baik	Baik
Yuliana Oktarina	Baik	Tidak Baik	Tidak Baik	Baik
Citra Akbar	Tidak Baik	Tidak Baik	Baik	Baik
Novi Putri	Baik	Baik	Baik	Baik
Maya Rambe	Baik	Baik	Tidak Baik	Baik
Riko Pratama	Tidak Baik	Baik	Baik	Tidak Baik
Vina Syah	Baik	Baik	Baik	Baik
Sari Fauzi	Baik	Baik	Baik	Baik
Rahma Anggraini	Baik	Baik	Baik	Tidak Baik
Karina Nugroho	Baik	Baik	Baik	Baik
Nadia Maulana	Tidak Baik	Baik	Baik	Baik
Bayu Pertiwi	Baik	Baik	Baik	Baik
Farhan Gunawan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Baik
Putri Santoso	Baik	Baik	Tidak Baik	Baik
Yuliana Irawan	Tidak Baik	Baik	Baik	Baik
Reza Fauzi	Baik	Baik	Baik	Baik
Citra Fikri	Baik	Baik	Baik	Baik
Rahma Lubis	Baik	Baik	Tidak Baik	Baik
Fajar Puspita	Baik	Baik	Tidak Baik	Tidak Baik
Arif Fikri	Baik	Baik	Tidak Baik	Baik
Rayhan Setiawan	Baik	Baik	Tidak Baik	Baik
Rahma Lestari	Baik	Baik	Baik	Baik
Arif Handayani	Baik	Baik	Baik	Baik
Ayu Hartono	Baik	Baik	Baik	Baik
Hendra Dewi	Baik	Baik	Baik	Baik
Sari Saputro	Tidak Baik	Tidak Baik	Baik	Baik
Surya Handayani	Tidak Baik	Tidak Baik	Baik	Baik
Yudi Sari	Tidak Baik	Baik	Baik	Baik
Yudi Syah	Tidak Baik	Tidak Baik	Baik	Tidak Baik
Citra Dewi	Baik	Baik	Tidak Baik	Tidak Baik
Andi Hartono	Baik	Tidak Baik	Baik	Baik
Novi Rambe	Tidak Baik	Baik	Baik	Baik
Rina Nasution	Baik	Tidak Baik	Baik	Tidak Baik
Fikri Nugroho	Baik	Baik	Baik	Tidak Baik
Fina Maharani	Baik	Baik	Baik	Baik
Dinda Fadilah	Baik	Baik	Baik	Baik
Bagas Fauzi	Tidak Baik	Baik	Baik	Baik
Rafi Nuraini	Baik	Baik	Baik	Baik
Lala Gunawan	Baik	Tidak Baik	Baik	Baik
Rani Kurniawan	Tidak Baik	Baik	Baik	Baik
Yuni Oktarina	Baik	Tidak Baik	Baik	Baik
Salsa Maulana	Baik	Baik	Baik	Tidak Baik

Pada tabel di atas disajikan data testing penelitian yang berjumlah 70 data responden pelanggan Teras Coffee Rantauprapat. Data testing merupakan data yang digunakan untuk menguji kemampuan model Decision Tree yang telah dibentuk

menggunakan data training. Setiap data terdiri atas empat variabel prediktor, yaitu kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan. Keempat variabel tersebut memiliki dua kategori penilaian, yaitu Baik dan Tidak Baik, yang digunakan sebagai dasar dalam proses klasifikasi. Berbeda dengan data training, data testing tidak digunakan untuk membentuk model, melainkan hanya digunakan sebagai data uji terhadap aturan keputusan yang telah dihasilkan. Penyusunan data dilakukan secara sistematis sehingga memudahkan proses pengujian pada setiap responden. Variasi kombinasi nilai pada setiap atribut memberikan peluang bagi model untuk menguji kemampuannya dalam mengenali pola yang telah dipelajari sebelumnya. Jumlah data testing sebanyak 70 responden memberikan ruang yang cukup untuk mengevaluasi performa model secara lebih menyeluruh. Semakin banyak data testing yang digunakan, semakin baik proses evaluasi terhadap kemampuan model dalam melakukan klasifikasi. Oleh karena itu, data pada tabel di atas menjadi bagian penting dalam mengukur tingkat keberhasilan metode Decision Tree. Data testing selanjutnya diproses menggunakan model Decision Tree yang telah terbentuk dari data training untuk menghasilkan prediksi tingkat kepuasan pelanggan. Setiap data diuji berdasarkan aturan-aturan keputusan yang terdapat pada pohon keputusan hingga diperoleh hasil klasifikasi sesuai karakteristik masing-masing responden. Proses pengujian ini bertujuan untuk mengetahui apakah model mampu mengklasifikasikan data baru dengan tepat. Hasil prediksi yang diperoleh kemudian dibandingkan dengan kondisi sebenarnya untuk mengukur tingkat akurasi model. Semakin banyak data yang berhasil diklasifikasikan dengan benar, maka semakin baik performa Decision Tree dalam memprediksi tingkat kepuasan pelanggan. Data testing juga berfungsi sebagai dasar dalam penyusunan confusion matrix dan perhitungan nilai evaluasi model, seperti accuracy, precision, recall, dan F1-score. Penggunaan 70 data testing memberikan hasil evaluasi yang lebih representatif karena jumlah data yang diuji lebih banyak dibandingkan data training. Variasi nilai pada setiap atribut juga membantu menguji konsistensi model terhadap berbagai kombinasi data yang berbeda. Hasil pengujian tersebut nantinya menjadi dasar untuk menilai kelayakan model dalam memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat. Dengan demikian, data testing pada tabel di atas memiliki peranan penting dalam memastikan bahwa model Decision Tree yang dibangun mampu menghasilkan klasifikasi yang akurat, konsisten, dan dapat diterapkan pada data baru.

Perancangan Model Klasifikasi

Perancangan model klasifikasi merupakan tahap yang dilakukan untuk membangun model prediksi tingkat kepuasan pelanggan berdasarkan data yang telah melalui proses pembersihan. Pada tahap ini, data yang tersedia diolah menggunakan algoritma Naïve Bayes Classifier dan Decision Tree untuk menghasilkan model yang mampu mengelompokkan pelanggan ke dalam kategori puas atau tidak puas. Proses perancangan model melibatkan penggunaan atribut seperti kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan sebagai variabel prediktor, sedangkan kepuasan pelanggan digunakan sebagai variabel target. Model yang terbentuk kemudian digunakan untuk mempelajari pola hubungan antarvariabel sehingga dapat menghasilkan prediksi terhadap data baru. Dengan adanya model klasifikasi ini, penelitian dapat mengetahui kemampuan masing-masing algoritma dalam mengidentifikasi tingkat kepuasan pelanggan secara tepat dan akurat.



Gambar 2. Perancangan Model Klasifikasi

Gambar di atas menunjukkan perancangan model klasifikasi yang digunakan dalam penelitian untuk memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat menggunakan aplikasi Orange Data Mining. Proses diawali dengan memasukkan data training melalui widget File, kemudian data tersebut digunakan sebagai masukan bagi algoritma Naïve Bayes dan Decision Tree (Tree) untuk membentuk model klasifikasi. Selain ditampilkan pada widget Data Table untuk melihat isi data yang digunakan, data training juga diproses oleh kedua algoritma guna mempelajari pola hubungan antara variabel kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan terhadap variabel target yaitu kepuasan pelanggan. Tahap ini bertujuan menghasilkan model yang mampu mengenali karakteristik pelanggan berdasarkan data yang telah tersedia.

Selanjutnya, data testing dimasukkan melalui *widget File* (1) dan dikirim ke widget Predictions bersama model yang telah dibentuk oleh algoritma *Naïve Bayes* dan Decision Tree. Pada tahap ini, sistem akan menghasilkan prediksi tingkat kepuasan pelanggan berdasarkan pola yang telah dipelajari sebelumnya oleh masing-masing algoritma. Hasil prediksi kemudian ditampilkan pada widget Data Table (2) sehingga dapat dianalisis dan dibandingkan sebelum disimpan melalui widget Save Data. Rancangan model klasifikasi ini digunakan untuk mengetahui kemampuan kedua algoritma dalam melakukan prediksi serta membandingkan hasil yang diperoleh guna menentukan metode yang memiliki performa terbaik dalam mengklasifikasikan tingkat kepuasan pelanggan.

Hasil Klasifikasi

Hasil klasifikasi merupakan keluaran yang dihasilkan oleh algoritma Naïve Bayes Classifier dan Decision Tree setelah melakukan proses prediksi terhadap data pelanggan. Hasil ini menunjukkan kategori tingkat kepuasan pelanggan, yaitu "Puas" atau "Tidak Puas", berdasarkan atribut kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan yang dimiliki setiap pelanggan. Melalui hasil klasifikasi tersebut, dapat diketahui pola hubungan antara variabel yang digunakan serta kemampuan masing-masing algoritma dalam mengelompokkan data secara tepat. Hasil yang diperoleh selanjutnya digunakan sebagai dasar untuk melakukan evaluasi dan membandingkan performa kedua metode klasifikasi yang diterapkan dalam penelitian.

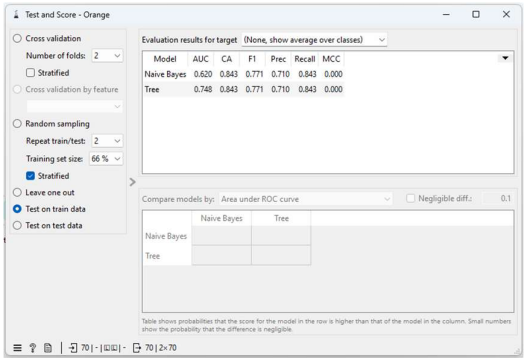
Tabel 2. Hasil Klasifikasi

Metode	Total Data	Puas	Tidak Puas
Naïve Bayes	29	54	16
Decision Tree	29	59	11

Berdasarkan tabel hasil klasifikasi di atas, terlihat adanya perbedaan hasil prediksi yang dihasilkan oleh metode Naïve Bayes dan Decision Tree terhadap data penelitian. Kedua metode menggunakan jumlah data yang sama, yaitu sebanyak 70 data testing, sehingga hasil yang diperoleh dapat dibandingkan secara objektif. Metode Naïve Bayes menghasilkan 54 data yang diklasifikasikan ke dalam kategori Puas dan 16 data ke dalam kategori Tidak Puas. Sementara itu, metode Decision Tree menghasilkan 59 data dengan kategori Puas dan 11 data dengan kategori Tidak Puas. Perbedaan jumlah hasil klasifikasi tersebut menunjukkan bahwa setiap algoritma memiliki mekanisme yang berbeda dalam mengenali pola hubungan antarvariabel. Naïve Bayes melakukan klasifikasi berdasarkan pendekatan probabilitas, sedangkan Decision Tree menggunakan aturan keputusan yang dibentuk dari atribut dengan nilai information gain tertinggi. Oleh karena itu, meskipun menggunakan data yang sama, hasil klasifikasi kedua metode tidak sepenuhnya identik. Perbedaan tersebut merupakan hal yang wajar dalam penelitian Machine Learning karena setiap algoritma memiliki karakteristik dan cara kerja yang berbeda. Hasil klasifikasi ini menjadi dasar untuk melakukan tahap evaluasi performa masing-masing metode pada proses selanjutnya. Hasil klasifikasi tersebut menunjukkan bahwa metode Decision Tree menghasilkan jumlah prediksi kategori Puas yang lebih banyak dibandingkan metode Naïve Bayes, yaitu 59 data, sedangkan Naïve Bayes menghasilkan 54 data. Sebaliknya, jumlah prediksi kategori Tidak Puas pada Decision Tree lebih sedikit, yaitu 11 data, sedangkan Naïve Bayes menghasilkan 16 data. Perbedaan hasil ini mengindikasikan bahwa Decision Tree membentuk aturan klasifikasi yang berbeda berdasarkan struktur pohon keputusan yang dihasilkan selama proses pelatihan. Sementara itu, Naïve Bayes menentukan hasil prediksi berdasarkan nilai probabilitas dari setiap atribut terhadap kelas yang tersedia. Kedua metode sama-sama mampu mengelompokkan data ke dalam kategori kepuasan pelanggan, namun distribusi hasil klasifikasinya berbeda. Hasil tersebut memberikan gambaran bahwa setiap metode memiliki keunggulan dalam mengenali karakteristik data yang digunakan. Selanjutnya, hasil klasifikasi ini akan dievaluasi menggunakan pengukuran performa seperti accuracy, precision, recall, dan F1-score untuk mengetahui metode yang memberikan tingkat prediksi terbaik. Dengan demikian, perbandingan hasil klasifikasi ini menjadi dasar dalam menentukan algoritma yang paling sesuai untuk memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat.

Hasil Evaluasi

Hasil evaluasi merupakan tahap yang dilakukan untuk menilai dan mengukur kinerja algoritma *Naïve Bayes Classifier* dan *Decision Tree* dalam melakukan klasifikasi tingkat kepuasan pelanggan. Evaluasi dilakukan dengan membandingkan hasil prediksi yang dihasilkan model dengan data aktual sehingga dapat diketahui tingkat ketepatan dan keandalan masing-masing algoritma. Melalui proses ini, peneliti dapat mengetahui seberapa baik model dalam mengklasifikasikan data pelanggan ke dalam kategori puas dan tidak puas. Hasil evaluasi juga digunakan sebagai dasar untuk membandingkan performa kedua metode klasifikasi serta menentukan algoritma yang paling sesuai untuk digunakan dalam memprediksi tingkat kepuasan pelanggan pada Teras Coffee Rantauprapat.



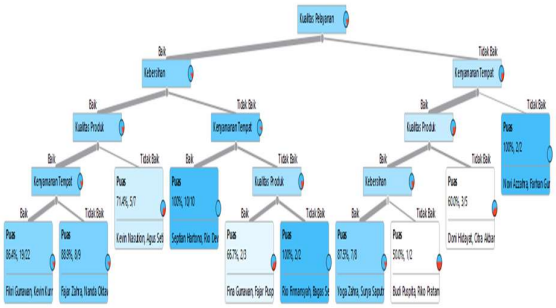
Gambar 3. Performance Metode Decision Tree

Gambar di atas menunjukkan hasil Test and Score menggunakan aplikasi Orange untuk membandingkan performa algoritma Naïve Bayes dan Decision Tree dalam mengklasifikasikan tingkat kepuasan pelanggan Teras Coffee Rantauprapat. Pengujian dilakukan menggunakan opsi Test on train data, sehingga model dievaluasi menggunakan data yang sama dengan data pelatihan. Hasil evaluasi ditampilkan berdasarkan beberapa metrik, yaitu AUC (Area Under Curve), CA (Classification Accuracy), F1-Score, Precision, Recall, dan MCC (Matthews Correlation Coefficient). Berdasarkan hasil tersebut, algoritma Naïve Bayes memperoleh nilai AUC sebesar 0,620, sedangkan Decision Tree memperoleh nilai AUC sebesar 0,748. Nilai AUC yang lebih tinggi pada Decision Tree menunjukkan bahwa algoritma tersebut memiliki kemampuan yang lebih baik dalam membedakan kelas Puas dan Tidak Puas dibandingkan Naïve Bayes. Sementara itu, kedua metode memperoleh nilai Classification Accuracy (CA) sebesar 0,843, yang menunjukkan bahwa keduanya mampu mengklasifikasikan data dengan tingkat akurasi sebesar 84,3%. Selain itu, nilai F1-Score sebesar 0,771, Precision sebesar 0,710, dan Recall sebesar 0,843 juga diperoleh oleh kedua algoritma, sehingga menunjukkan performa klasifikasi yang relatif seimbang pada metrik tersebut.

Hasil pengujian juga memperlihatkan bahwa meskipun nilai akurasi kedua metode sama, terdapat perbedaan pada nilai AUC yang cukup signifikan. Nilai AUC sebesar 0,748 pada Decision Tree menunjukkan bahwa model memiliki kemampuan yang lebih baik dalam membedakan data antar kelas dibandingkan Naïve Bayes yang hanya memperoleh nilai 0,620. Semakin tinggi nilai AUC, maka semakin baik kemampuan model dalam melakukan proses klasifikasi secara keseluruhan. Sementara itu, nilai Precision sebesar 0,710 menunjukkan bahwa sekitar 71% dari seluruh data yang diprediksi sebagai kategori tertentu merupakan prediksi yang benar. Nilai Recall sebesar 0,843 menunjukkan bahwa model berhasil mengenali sekitar 84,3% data yang sebenarnya termasuk dalam kelas target. Nilai F1-Score sebesar 0,771 merupakan hasil keseimbangan antara nilai Precision dan Recall sehingga menggambarkan kualitas klasifikasi secara menyeluruh. Pada metrik Matthews Correlation Coefficient (MCC), kedua metode memperoleh nilai 0,000. Nilai tersebut menunjukkan bahwa koefisien korelasi yang dihitung oleh Orange tidak memberikan informasi tambahan yang signifikan terhadap performa model pada konfigurasi pengujian yang digunakan. Oleh karena itu, perbandingan performa lebih tepat difokuskan pada nilai AUC, CA, F1-Score, Precision, dan Recall.

Berdasarkan hasil Test and Score tersebut dapat disimpulkan bahwa algoritma Decision Tree memberikan performa yang sedikit lebih baik dibandingkan algoritma Naïve Bayes, terutama berdasarkan nilai AUC yang lebih tinggi. Walaupun kedua metode menghasilkan nilai akurasi yang sama, kemampuan Decision Tree dalam membedakan

antar kelas menunjukkan hasil yang lebih optimal. Hal ini mengindikasikan bahwa struktur pohon keputusan yang dibentuk oleh Decision Tree mampu menangkap pola hubungan antar atribut dengan lebih baik dibandingkan pendekatan probabilitas yang digunakan oleh Naïve Bayes. Kesamaan nilai pada metrik CA, F1-Score, Precision, dan Recall menunjukkan bahwa kedua algoritma memiliki tingkat keberhasilan klasifikasi yang relatif sebanding pada data penelitian. Namun, keunggulan Decision Tree pada nilai AUC menjadi indikator bahwa algoritma tersebut lebih stabil dalam proses pemisahan kelas. Hasil pengujian ini mendukung proses pemilihan metode yang paling sesuai untuk memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat. Dengan demikian, berdasarkan hasil evaluasi menggunakan fitur Test and Score pada Orange, algoritma Decision Tree dapat dinilai memiliki performa yang lebih baik dibandingkan algoritma Naïve Bayes, meskipun selisih performanya tidak terlalu besar.



Gambar 4. Pohon Keputusan Metode Decision Tree

Pada gambar di atas ditampilkan hasil visualisasi Decision Tree yang diperoleh dari proses pengolahan data menggunakan aplikasi Orange. Pohon keputusan tersebut dibentuk berdasarkan hasil perhitungan nilai entropy dan information gain pada setiap atribut, sehingga menghasilkan struktur keputusan yang digunakan untuk mengklasifikasikan tingkat kepuasan pelanggan. Atribut yang berada pada bagian paling atas atau root node adalah Kualitas Pelayanan, yang menunjukkan bahwa atribut tersebut memiliki kemampuan terbaik dalam membedakan data dibandingkan atribut lainnya. Dari node akar tersebut, data kemudian dibagi menjadi dua cabang utama, yaitu kategori Baik dan Tidak Baik. Selanjutnya, setiap cabang akan kembali dipisahkan berdasarkan atribut lain seperti Kebersihan, Kualitas Produk, dan Kenyamanan Tempat hingga mencapai node akhir atau leaf node. Setiap leaf node menunjukkan hasil klasifikasi berupa kategori Puas beserta persentase probabilitas dan jumlah data yang berada pada node tersebut. Struktur pohon keputusan ini memberikan gambaran yang jelas mengenai proses pengambilan keputusan yang dilakukan oleh algoritma Decision Tree. Selain itu, visualisasi tersebut mempermudah pengguna dalam memahami hubungan antarvariabel yang memengaruhi tingkat kepuasan pelanggan. Dengan demikian, pohon keputusan tidak hanya berfungsi sebagai model klasifikasi, tetapi juga sebagai media interpretasi hasil analisis yang mudah dipahami.

Berdasarkan struktur pohon keputusan tersebut terlihat bahwa atribut Kualitas Pelayanan menjadi faktor utama yang menentukan proses klasifikasi tingkat kepuasan pelanggan. Apabila pelanggan memberikan penilaian Baik terhadap kualitas pelayanan, maka proses klasifikasi akan dilanjutkan dengan melihat atribut Kebersihan sebagai atribut berikutnya. Setelah itu, proses pengambilan keputusan kembali mempertimbangkan atribut Kualitas Produk dan Kenyamanan Tempat hingga diperoleh hasil akhir berupa kategori Puas. Beberapa cabang menunjukkan tingkat probabilitas kepuasan yang sangat tinggi, seperti 86,4%, 88,9%, bahkan 100%, yang menandakan

bahwa kombinasi atribut tertentu sangat berpengaruh terhadap kepuasan pelanggan. Sebaliknya, apabila kualitas pelayanan memperoleh penilaian Tidak Baik, maka proses klasifikasi lebih banyak mengarah pada atribut Kenyamanan Tempat sebelum menghasilkan keputusan akhir. Hal tersebut menunjukkan bahwa setiap atribut memiliki peran yang berbeda sesuai dengan kondisi data yang dianalisis. Percabangan yang terbentuk juga memperlihatkan bahwa tidak semua pelanggan memiliki pola penilaian yang sama. Oleh karena itu, Decision Tree mampu menghasilkan aturan keputusan yang fleksibel berdasarkan karakteristik masing-masing data. Kejelasan struktur pohon ini menjadi salah satu keunggulan metode Decision Tree dibandingkan metode klasifikasi lainnya karena aturan yang dihasilkan dapat dipahami secara langsung.

Hasil visualisasi tersebut juga memperlihatkan bahwa sebagian besar leaf node menghasilkan klasifikasi Puas, yang menunjukkan bahwa mayoritas responden memberikan penilaian positif terhadap layanan yang diberikan oleh Teras Coffee Rantauprapat. Beberapa node bahkan menghasilkan tingkat kepastian sebesar 100%, yang berarti seluruh data pada cabang tersebut berada dalam kategori yang sama sehingga tidak lagi memerlukan proses pemisahan lanjutan. Selain itu, terdapat beberapa leaf node yang memiliki tingkat kepastian sekitar 50% hingga 71,4%, yang menunjukkan bahwa masih terdapat variasi data pada cabang tersebut. Kondisi ini menunjukkan bahwa tidak semua atribut memiliki kemampuan yang sama dalam membedakan tingkat kepuasan pelanggan. Semakin tinggi nilai persentase pada suatu leaf node, maka semakin tinggi pula tingkat keyakinan model dalam memberikan hasil klasifikasi. Hasil tersebut menunjukkan bahwa algoritma Decision Tree mampu membangun aturan klasifikasi berdasarkan pola hubungan antarvariabel yang terdapat pada data penelitian. Aturan-aturan tersebut dapat dijadikan sebagai dasar dalam memahami faktor-faktor yang memengaruhi kepuasan pelanggan secara lebih rinci. Dengan adanya visualisasi pohon keputusan ini, proses interpretasi hasil penelitian menjadi lebih mudah karena setiap jalur keputusan dapat ditelusuri mulai dari node akar hingga node akhir. Oleh sebab itu, hasil Decision Tree tidak hanya memberikan prediksi, tetapi juga memberikan penjelasan mengenai alasan terbentuknya setiap keputusan klasifikasi.

Berdasarkan keseluruhan struktur pohon keputusan, dapat diketahui bahwa kombinasi antara Kualitas Pelayanan, Kebersihan, Kualitas Produk, dan Kenyamanan Tempat memiliki pengaruh yang berbeda-beda terhadap tingkat kepuasan pelanggan. Urutan atribut yang muncul pada setiap percabangan menunjukkan tingkat kepentingan masing-masing atribut dalam proses klasifikasi. Atribut yang muncul lebih awal memiliki pengaruh yang lebih besar dibandingkan atribut yang berada pada percabangan berikutnya. Hal ini menunjukkan bahwa Decision Tree mampu memilih atribut terbaik secara otomatis berdasarkan hasil perhitungan information gain. Model yang dihasilkan juga memberikan aturan keputusan yang sederhana sehingga mudah dipahami oleh pengguna maupun pihak manajemen. Informasi tersebut dapat dimanfaatkan sebagai bahan evaluasi dalam meningkatkan kualitas pelayanan kepada pelanggan. Pihak pengelola dapat memprioritaskan perbaikan pada atribut yang berada di bagian awal pohon keputusan karena atribut tersebut memiliki kontribusi terbesar terhadap kepuasan pelanggan. Selain itu, visualisasi ini juga memudahkan proses pengambilan keputusan karena hubungan antarvariabel dapat diamati secara langsung. Dengan demikian, hasil Decision Tree memberikan informasi yang tidak hanya akurat dalam proses klasifikasi, tetapi juga mudah diinterpretasikan sehingga dapat mendukung pengambilan keputusan yang lebih efektif di Teras Coffee Rantauprapat.

5. Kesimpulan

Penelitian ini berhasil menerapkan algoritma Naïve Bayes Classifier dan Decision Tree untuk memprediksi tingkat kepuasan pelanggan Teras Coffee Rantauprapat berdasarkan variabel kualitas pelayanan, kualitas produk, kenyamanan tempat, dan kebersihan. Proses penelitian diawali dengan pengumpulan data melalui penyebaran kuesioner, kemudian dilanjutkan dengan tahap pembagian data, transformasi data, pengolahan menggunakan Microsoft Excel, serta implementasi model menggunakan aplikasi Orange. Hasil klasifikasi menunjukkan bahwa kedua metode mampu mengelompokkan data pelanggan ke dalam kategori Puas dan Tidak Puas dengan tingkat akurasi yang baik. Berdasarkan hasil evaluasi Test and Score, kedua metode memperoleh nilai Classification Accuracy (CA) sebesar 84,3%, sedangkan algoritma Decision Tree memiliki nilai AUC sebesar 0,748, lebih tinggi dibandingkan Naïve Bayes yang memperoleh nilai AUC sebesar 0,620. Hasil tersebut menunjukkan bahwa kedua metode layak digunakan dalam memprediksi tingkat kepuasan pelanggan, namun Decision Tree memiliki kemampuan yang lebih baik dalam membedakan setiap kelas data.

Berdasarkan hasil visualisasi pohon keputusan, diketahui bahwa atribut Kualitas Pelayanan menjadi faktor utama yang memengaruhi tingkat kepuasan pelanggan, kemudian diikuti oleh atribut Kebersihan, Kualitas Produk, dan Kenyamanan Tempat pada percabangan berikutnya. Informasi tersebut menunjukkan bahwa kualitas pelayanan memiliki kontribusi paling besar dalam membentuk keputusan klasifikasi pelanggan. Hasil penelitian ini diharapkan dapat menjadi bahan evaluasi bagi pihak Teras Coffee Rantauprapat dalam meningkatkan kualitas pelayanan, fasilitas, dan kenyamanan yang diberikan kepada pelanggan. Selain itu, penerapan metode Machine Learning memberikan pendekatan yang lebih objektif dibandingkan metode penilaian konvensional karena mampu menghasilkan prediksi berdasarkan pola data yang dimiliki. Dengan demikian, hasil penelitian ini dapat mendukung proses pengambilan keputusan yang lebih tepat, efektif, dan berbasis data dalam upaya meningkatkan kepuasan pelanggan.

6. Daftar Pustaka

Aditya, N., Munthe, I. R., & Sihombing, V. (2024). *A Comparative Analysis of Machine Learning Algorithms for Predicting Paddy Production*. *8*(2), 1200–1207.

Adjani, K., Fauzia, F. A., & Juliane, C. (2023). Comparison of K-N Earest 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

Anam, K., Nurhakim, B., & Juliane, C. (2022). Komparasi Algoritma Klasifikasi Data Mining Menggunakan Optimize Selection untuk Peminatan Program Studi. *Building of Informatics, Technology and Science (BITS)*, *4*(2), 606–613. https://doi.org/10.47065/bits.v4i2.2160

Anwar, B., Ambiyar, A., & Fadhilah, F. (2023). Application of the FP-Growth Method to Determine Drug Sales Patterns. *Sinkron*, *8*(1), 405–414. https://doi.org/10.33395/sinkron.v8i1.12004

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

Ayuningtyas, N., Nining, R., & Basysyar, F. M. (2022). Penerapan Data Mining pada Penjualan Produk MS Glow Menggunakan Metode Klasifikasi untuk Strategi Pemasaran. *Jurnal Accounting Information System (AIMS*, *5*(2), 156–166. Retrieved from https://jurnal.masoemuniversity.ac.id/index.php/aims

Diana Dewi, D., Qisthi, N., Lestari, S. S. S., & Putri, Z. H. S. (2023). Perbandingan Metode Neural Network Dan Support Vector Machine Dalam Klasifikasi Diagnosa Penyakit Diabetes. *Cerdika: Jurnal Ilmiah Indonesia*, *3*(09), 828–839. https://doi.org/10.59141/cerdika.v3i09.662

Diansyah, S. (2022). *Jurnal Sistim Informasi dan Teknologi Klasifikasi Tingkat Kepuasan Pengguna dengan Menggunakan Metode K-Nearest Neighbour (KNN)*. *4*, 1–3. https://doi.org/10.37034/jsisfotek.v4i1.114

Hasri, C. F., & Alita, D. (2022). Penerapan Metode Naã Ve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, *3*(2), 145–160. https://doi.org/10.33365/jatika.v3i2.2026

Juliadi, R. N., & Puspitarani, Y. (2022). Supervised Model for Sentiment Analysis Based on Hotel Review Clusters using RapidMiner. *SinkrOn*, *7*(3), 1059–1066. https://doi.org/10.33395/sinkron.v7i3.11564

Melyani, S., & Harahap, S. Z. (2024). *Prediction of Stunting in Toddlers Combining the Naive Bayes Method and the C4 . 5 Algorithm*. *8*(2), 1160–1168.

Nugraha, A. B., & Romadhony, A. (2023). Identification of 10 Regional Indonesian Languages Using Machine Learning. *Sinkron*, *8*(4), 2203–2214. https://doi.org/10.33395/sinkron.v8i4.12989

Pratama, H. A., Yanris, G. J., Nirmala, M., & Hasibuan, S. (2023). *Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels*. *8*(3), 1832–1851.

Ramadhon, R. N., Ogi, A., & Agung, A. P. (2024). *Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank*. *3*, 1860–1874.

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

Wahyudi, A., Ovelia Tampubolon, S., afrilia Putri, N., Ghassa, A., Rasywir, E., & Kisbianty, D. (2022). Penerapan Data Mining Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan Terhadap INDIHOME. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, *2*(2), 240–247. https://doi.org/10.33998/jakakom.2022.2.2.111

Wijaya, E. B., Dharma, A., Heyneker, D., & Vanness, J. (2023). Comparison of the K-Means Algorithm and C4.5 Against Sales Data. *SinkrOn*, *8*(2), 741–751. https://doi.org/10.33395/sinkron.v8i2.12224

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