



## ***Implementing the C4.5 Algorithm for Customer Satisfaction Classification***

### **Implementasi Algoritma C4.5 untuk Klasifikasi Kepuasan Pelanggan**

**Listina Nadhia Ningsih<sup>1\*</sup>, Riska Septiani<sup>2</sup>,  
Angga Pramadjaya<sup>3</sup>, Seli Nuralisah<sup>4</sup>**

<sup>1,2,3,4</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer,  
Universitas Pamulang, Indonesia

E-Mail: <sup>1</sup>dosen02875@unpam.ac.id, <sup>2</sup>dosen02912@unpam.ac.id,  
<sup>3</sup>dosen10029@unpam.ac.id, <sup>4</sup>selinuralisah@gmail.com.

Received Jul 30th 2025; Revised Oct 20th 2025; Accepted Oct 31th 2025; Available Online Nov 4th 2025

Corresponding Author: Listina Nadhia Ningsih

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

#### **Abstract**

*The development of the culinary industry in Indonesia shows an increasingly competitive trend, driven by the increasing number of business actors and dynamic changes in consumer preferences. In this situation, customer satisfaction is a key factor that determines the sustainability and growth of businesses, including traditional culinary businesses such as the Ayam Bekakak Restaurant. However, conventional customer satisfaction data management is often subjective and does not produce information that can be processed systematically. This makes data-driven decision-making difficult. Therefore, an analytical method is needed that can process customer satisfaction data more accurately and objectively. One solution that can be implemented is the use of data mining methods, specifically the C4.5 algorithm. The results of the study showed that the C4.5 algorithm was able to classify customer satisfaction levels with an accuracy of 96.15%, with a precision value of 100% and a recall of 95.2%. This value indicates that the model can predict customer satisfaction with a very low error rate. The confusion matrix results showed only one misclassified data item, where a customer who was actually "satisfied" was detected as "dissatisfied." This indicates that the model has high generalization ability and minimal classification errors. Compliance of data characteristics with the C4.5 algorithm.*

**Keyword:** C4.5 Algorithm, Classification, Implementation, Satisfaction, Customers.

#### **Abstrak**

Perkembangan industri kuliner di Indonesia menunjukkan tren yang semakin kompetitif, didorong oleh meningkatnya jumlah pelaku usaha dan perubahan preferensi konsumen yang dinamis. Dalam situasi ini, kepuasan pelanggan menjadi faktor kunci yang sangat menentukan keberlangsungan dan pertumbuhan bisnis, termasuk bagi usaha kuliner tradisional seperti *Restaurant Ayam Bekakak*. Namun, pengelolaan data kepuasan pelanggan yang selama ini dilakukan secara konvensional sering kali bersifat subjektif dan tidak menghasilkan informasi yang dapat diolah secara sistematis. Hal ini menyebabkan sulitnya dalam pengambilan keputusan yang berbasis data (*data-driven decision making*). Oleh karena itu, dibutuhkan metode analisis yang mampu mengolah data kepuasan pelanggan secara lebih akurat dan objektif. Salah satu solusi yang dapat diterapkan adalah dengan menggunakan metode *data mining*, khususnya algoritma C4.5. Berdasarkan hasil penelitian menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan tingkat kepuasan pelanggan dengan akurasi sebesar 96,15%, dengan nilai *precision* 100% dan *recall* 95,2%. Nilai ini menunjukkan bahwa model dapat memprediksi kepuasan pelanggan dengan tingkat kesalahan yang sangat kecil. Dari hasil *confusion matrix*, hanya satu data yang salah klasifikasi, di mana pelanggan yang sebenarnya "puas" terdeteksi sebagai "tidak puas". Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang tinggi dan kesalahan klasifikasi yang minim. Kesesuaian karakteristik data dengan algoritma C4.5.

**Kata Kunci:** Algoritma C4.5, Implementasi, Klasifikasi, Kepuasan, Pelanggan.

#### **1. PENDAHULUAN**

Perkembangan industri kuliner di Indonesia menunjukkan tren yang semakin kompetitif, didorong oleh meningkatnya jumlah pelaku usaha dan perubahan preferensi konsumen yang dinamis[1]. Dalam situasi ini, kepuasan pelanggan menjadi faktor kunci yang sangat menentukan keberlangsungan dan pertumbuhan bisnis, termasuk bagi usaha kuliner tradisional seperti *Restaurant Ayam Bekakak*. Kepuasan pelanggan tidak

hanya mencerminkan keberhasilan layanan yang diberikan, tetapi juga berperan penting dalam membentuk loyalitas pelanggan dan citra positif usaha di mata publik[2]. Di era persaingan bisnis kuliner yang semakin kompetitif, kepuasan pelanggan menjadi salah satu indikator penting dalam mempertahankan eksistensi dan meningkatkan loyalitas konsumen. *Restaurant* sebagai penyedia jasa makanan tidak hanya dituntut menyajikan makanan berkualitas, tetapi juga memberikan pelayanan yang mampu memenuhi ekspektasi pelanggan[3]. Kepuasan pelanggan merupakan salah satu indikator kunci dalam menilai keberhasilan suatu perusahaan, khususnya yang bergerak di bidang jasa atau pelayanan. Tingkat kepuasan pelanggan dapat memengaruhi loyalitas, citra perusahaan, serta keberlangsungan bisnis dalam jangka panjang[4]. Kepuasan pelanggan merupakan elemen fundamental dalam menentukan keberhasilan dan keberlangsungan suatu bisnis. Pelanggan yang merasa puas cenderung memiliki loyalitas tinggi, memberikan ulasan positif, dan menjadi promotor bagi produk atau layanan yang digunakan[5]. Oleh karena itu, perusahaan dituntut untuk terus memantau dan mengevaluasi tingkat kepuasan pelanggan guna menjaga daya saing dan meningkatkan kualitas layanan[6]. Dalam era digital dan perkembangan teknologi informasi saat ini, perusahaan memiliki akses terhadap beragam data pelanggan yang sangat besar (*big data*), baik berupa survei, ulasan, maupun umpan balik secara langsung[7]. Namun, pengolahan data tersebut secara manual tidak efektif dan berisiko menimbulkan bias. Diperlukan pendekatan yang lebih sistematis dan cerdas untuk mengekstraksi informasi bermakna dari data yang tersedia[8]. Oleh karena itu, penting bagi pihak *restaurant* untuk secara rutin mengevaluasi dan mengklasifikasikan tingkat kepuasan pelanggan guna memperoleh umpan balik yang konstruktif serta melakukan perbaikan terhadap aspek-aspek layanan yang kurang optimal[9].

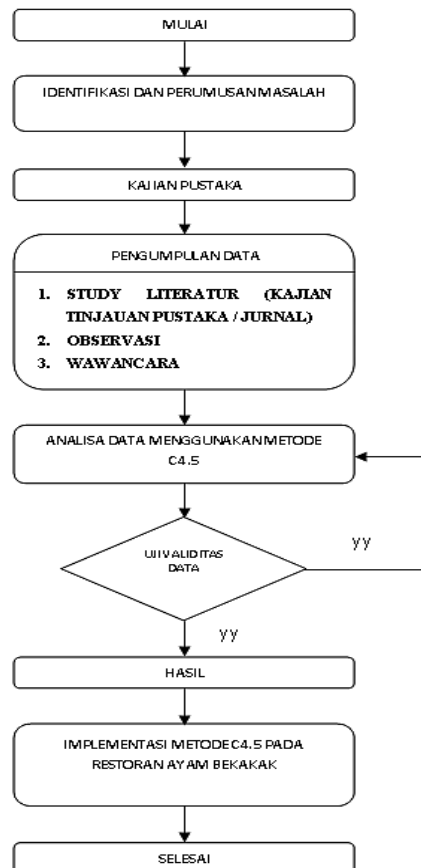
Dalam praktiknya, data kepuasan pelanggan sering kali bersifat kompleks dan tersebar dalam jumlah besar, sehingga analisis manual menjadi tidak efisien dan rawan terhadap subjektivitas[10]. Untuk menjawab tantangan ini, pendekatan berbasis *data mining* menjadi solusi yang efektif. *Data mining* memungkinkan perusahaan untuk menggali pola tersembunyi dari data yang ada dan mengubahnya menjadi informasi yang berguna dalam pengambilan keputusan strategis[11]. *Data mining* yang bisa digunakan ialah menggunakan metode klasifikasi yang digunakan untuk membentuk pohon keputusan (*decision tree*). Algoritma ini merupakan pengembangan dari ID3 (*Iterative Dichotomiser 3*). Tujuannya adalah untuk mengklasifikasikan data berdasarkan atribut-atribut tertentu dan memprediksi kelas atau kategori dari data baru. Dalam konteks kepuasan pelanggan, algoritma ini dapat membantu menentukan apakah pelanggan termasuk kategori “puas” atau “tidak puas” berdasarkan berbagai faktor. C4.5 dapat digunakan untuk memprediksi tingkat kepuasan pelanggan berdasarkan atribut-atribut seperti kualitas pelayanan, harga produk, promosi, respon layanan pelanggan, kenyamanan fasilitas, loyalitas pelanggan. Data ini biasanya diperoleh dari kuesioner atau survei pelanggan yang memiliki label target berupa Puas Dan Tidak puas. Atau bisa juga multi-level: “Sangat puas”, “Cukup puas”, “Tidak puas”. Aapun langkah-langkah dalam klasifikasi kepuasan pelanggan yaitu memilih atribut yang paling berpengaruh terhadap kepuasan pelanggan menggunakan nilai *Gain Ratio* (rasio antara *Information Gain* dan *Split Information*). Membentuk simpul akar (*root node*) Atribut dengan *Gain Ratio* tertinggi dijadikan *node* utama. Membagi data menjadi cabang berdasarkan nilai atribut tersebut. Mengulangi proses untuk setiap cabang hingga semua data terklasifikasi atau tidak ada atribut tersisa. Membangun aturan keputusan (*decision rules*) dari pohon yang terbentuk, misalnya: Jika Kualitas\_Pelayanan = Baik dan Harga = Terjangkau Maka Kepuasan = Puas.

Selain itu, penelitian yang dilakukan oleh sinaga (2021) menunjukkan bahwa klasifikasi menggunakan algoritma C4.5 untuk kepuasan pelanggan pada perusahaan air minum (PDAM) di Kabupaten Simalungun dengan 5 variabel yaitu fasilitas layanan, tarif, kelancaran air, citra korporasi, serta lokasi menunjukkan hasil akurasi sangat tinggi 98% untuk klasifikasi “Puas” vs “Tidak Puas”. Adaun relevansinya menunjukkan bahwa C4.5 telah berhasil digunakan dalam konteks layanan *public*/korporasi di Indonesia, serta menunjukkan variabel-faktor kepuasan pelanggan yang umum. Penelitian yang dilakukan oleh hermanto (2024) menunjukkan bahwa algoritma C4.5 untuk menganalisis data pelanggan yang merasa puas atau tidak puas dan memberikan relevansi dengan memberikan contoh penerapan pada ritel, menunjukkan fleksibilitas algoritma C4.5 dalam berbagai domain pelanggan. Penelitian yang dilakukan oleh Dewi (2025) bahwa aplikasi pada bank Syariah (Nasabah), dengan atribut seperti layanan, fasilitas, kemudahan akses, produk, dan kepercayaan. Mendapatkan hasil akurasi mencapai 96,31% memberikan relevansi dengan mengangkat sector jasa keuangan dan faktor kepercayaan sebagai variabel kepuasan bisa memperluas perspektif variabel . Penelitian tersebut relevan dengan penelitian ini yaitu penelitian ini tidak hanya menerapkan C4.5 untuk memprediksi kepuasan, tetapi juga menguji teknik penanganan data tidak seimbang, mengintegrasikan hasil ke CRM, dan melakukan evaluasi intervensi untuk membuktikan dampak praktis”.

Penelitian ini mengimplementasikan algoritma C4.5 dengan kombinasi variabel baru yang spesifik pada konteks pelanggan tertentu (misalnya pelanggan ritel, jasa, atau sektor pertanian/peternakan). Melakukan pengujian akurasi dan validasi model secara sistematis, misalnya dengan *cross validation* atau *confusion matrix*, sehingga hasilnya lebih kuat secara ilmiah. Dapat mengusulkan modifikasi atau pengembangan aturan keputusan yang lebih adaptif terhadap pola data lokal (lokasi, budaya, atau jenis usaha tertentu). Tidak hanya membangun model klasifikasi, tetapi juga mengintegrasikan hasil klasifikasi ke sistem *Customer Relationship Management* (CRM) agar hasilnya bisa langsung digunakan untuk strategi retensi

pelanggan. Menggunakan data riil pelanggan dari institusi atau mitra lokal (bukan data simulasi), sehingga model yang dihasilkan lebih representatif terhadap kondisi nyata di lapangan. Menghasilkan *dashboard* atau visualisasi pohon keputusan yang membantu pihak manajemen memahami hasil model tanpa latar belakang teknis mendalam.

## 2. METODE PENELITIAN



**Gambar 1.** Tahapan Penelitian

Pada Gambar 1 tahapan yang dilakukan dalam penelitian adalah identifikasi masalah dan perumusan masalah. Pada tahap pertama penelitian dimulai dengan mengidentifikasi serta merumuskan masalah pada permasalahan yang diteliti pada Restoran Ayam Bekakak. Adapun tahapan kajian Pustaka yang bisa dilakukan ialah dengan tahapan mencari teori-teori yang berkaitan dengan topik penelitian yang berasal dari sumber-sumber penelitian kepustakaan. Pengumpulan data yaitu teknik pengumpulan data yang digunakan pada penelitian ini antara lain dengan penelitian Lapangan (*field research*), dilakukan guna memperoleh data primer yang dibutuhkan dalam penelitian dengan cara pengamatan secara langsung di lapangan (*field research*) dilakukan dengan cara mengobservasi yaitu pengamatan yang dilakukan secara langsung terhadap aktivitas di Restoran Ayam Bekakak. Selain itu Wawancara yaitu melakukan tanya jawab secara langsung kepada pihak terkait dengan mengajukan pertanyaan-pertanyaan. Dengan menggunakan penelitian kepustakaan (*Library Research*), dilakukan untuk memperoleh data sekunder dengan melakukan penelaahan terhadap teori-teori yang berkaitan dengan topik penelitian yang berasal dari sumber-sumber penelitian kepustakaan. Sumber-sumber penelitian kepustakaan dapat diperoleh dari jurnal, buku, majalah, hasil-hasil penelitian yang terdahulu dan telah dipublikasikan (tesis dan disertasi) dan sumber lainnya (internet, surat kabar dan lain-lain) yang sesuai dengan topik penelitian.

Analisis data yang digunakan dalam tahap ini dilakukan pengkajian data yang diperoleh berdasarkan metode Algoritma C4.5. Selanjutnya Uji validitas data yaitu pada tahap ini data yang telah dikumpulkan menggunakan microsoft excel akan diuji menggunakan dengan metode Algoritma C4.5. Jika data yang diuji tidak valid maka data akan kembali dianalisis pada tahapan sebelumnya. Hasil yang diharapkan pada tahap ini *output* berupa hasil pengklasifikasian kepuasan pelanggan yang nantinya dapat digunakan untuk mengklasifikasi kepuasan pelanggan pada Restoran Ayam Bekakak. Adapun Implementasi yang bisa dilakukan pada tahap ini ialah Algoritma C4.5 akan diimplementasikan untuk pengklasifikasian kepuasan pelanggan Restoran Ayam Bekakak dengan cara data kepuasan pelanggan yang telah diambil akan diolah

dengan Algoritma C4.5 dan menghasilkan pohon keputusan sehingga bisa dilakukannya Klasifikasi Kepuasan Pelanggan

## 2.1 Klasifikasi

Klasifikasi adalah proses pengelompokkan / pengumpulan benda atau entitas yang sama, serta memisahkan benda atas entitas yang tidak sama[12]. Dalam pengertian secara umum bahwa klasifikasi ialah suatu kegiatan yang mengelompokkan benda yang memiliki beberapa ciri yang sama dan memisahkan benda yang tidak sama. Dalam kaitannya di dunia perpustakaan klasifikasi diartikan sebagai kegiatan pengelompokkan bahan pustaka berdasarkan ciri-ciri yang sama, misalnya pengarang, fisik, isi dsb. Pada dasarnya di perpustakaan dikenal ada 2 (dua) jenis kegiatan klasifikasi[13].

1. Klasifikasi Fundamental (*Fundamental Classification*) yaitu klasifikasi bahan pustaka berdasarkan subjek/isi buku, sebab pada dasarnya pemakai perpustakaan lebih banyak mencari informasi tentang subjek tertentu.
2. Klasifikasi Artifisial (*Artificial Classification*) yaitu klasifikasi bahan pustaka berdasarkan ciri-ciri yang ada pada bahan pustaka. Misalnya klasifikasi berdasarkan warna, ukuran dsb.

## 2.2. Analisis Menggunakan Algoritma C4.5

Untuk melakukan Klasifikasi Kepuasan pelanggan pada Restoran Ayam Bekakak, data yang digunakan adalah data hasil Kuisisioner yang dapat dilihat pada lampiran menggunakan Algoritma C4.5 dengan menggunakan persamaan sebagai berikut :

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$

$S$  ialah Himpunan kasus,  $A$  ialah Atribut  $n$  ialah Jumlah partisi atribut,  $|S_i|$  ialah Jumlah kasus pada partisi ke  $i$ .  $|S|$  : Jumlah kasus dalam  $S$ . Selanjutnya Menghitung *Gain*. Sebelum mendapatkan nilai *Gain* adalah dengan mencari nilai Entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut[14]. Rumus dasar dari *Gain* adalah sebagai berikut:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} * Entropy(S_i) \quad (2)$$

Hasil Klasifikasi menggunakan Algoritma C4.5 dilakukan dengan menghitung Data Kuisisioner yang telah di lampirkan. Berikut perhitungan Klasifikasi C4.5 sesuai dengan rumus yang telah diuraikan sebelumnya.

## 2.3 Literature Review

Penelitian yang dilakukan oleh Hermanto (2024) [15] Algoritma C4.5 banyak digunakan pada survei kepuasan (kuisisioner) dan data transaksi untuk mengklasifikasikan label seperti puas / cukup puas / tidak puas atau *satisfied* / *dissatisfied*. Contoh: penelitian pada *e-commerce*, *call center*, lembaga pelatihan, dan *multifinance*. Penelitian yang dilakukan oleh Steffany (2023) [16] menghasilkan banyak studi lokal/akademik melaporkan akurasi tinggi pada studi pada pengguna marketplace (Shopee) melaporkan akurasi ~94,17%; studi toko *online*/ritel lain melaporkan akurasi 91–94%; studi lain melaporkan akurasi sekitar 80–87% tergantung kualitas data dan *preprocessing*. Ini menunjukkan C4.5 sering kompetitif untuk *dataset* survei kepuasan. Penelitian yang dilakukan oleh Apriliyawan (2020) menghasilkan studi komparatif sering membandingkan C4.5 dengan Naïve Bayes, SVM, KNN Adapun hasil campuran: beberapa studi menemukan C4.5 memberikan akurasi lebih tinggi daripada Naïve Bayes pada *dataset* tertentu; pada *dataset* lain SVM/ KNN bisa lebih baik hasil tergantung pada ukuran data, distribusi fitur, dan fitur numerik vs kategorikal. Oleh karena itu banyak studi menyarankan uji beberapa model dan gunakan *cross-validation* [17]. Berdasarkan sintesis literatur dan temuan empiris, dapat disimpulkan bahwa algoritma C4.5 merupakan metode yang efektif dan efisien dalam mengklasifikasikan tingkat kepuasan pelanggan. Model ini unggul dalam hal interpretabilitas, kemudahan implementasi, serta kemampuannya mengidentifikasi atribut paling berpengaruh terhadap kepuasan pelanggan. Namun, untuk memperoleh hasil yang lebih optimal, penelitian harus memperhatikan proses *preprocessing* data, keseimbangan kelas, dan validasi model secara menyeluruh. Dengan demikian, penelitian lanjutan mengenai “Implementasi Algoritma C4.5 untuk Klasifikasi Kepuasan Pelanggan” memiliki urgensi yang tinggi, terutama dalam mendukung pengambilan keputusan berbasis data (*data-driven decision making*) untuk meningkatkan kualitas pelayanan dan loyalitas pelanggan.

## 3. HASIL DAN PEMBAHASAN

### 3.1. Perhitungan Algoritma C4.5

Proses perhitungan dengan menggunakan Algoritma diawali dengan. perhitungan *Entropy* dan *Gain* untuk menentukan akar (*root*) dari pohon keputusan, sampai terbentuk pohon keputusan. Data Kepuasan

Pelanggan yang di teliti merupakan data hasil dari pengumpulan kuisioner kepuasan pelanggan yang dilakukan di Restoran Ayam Bekakak selama satu bulan yaitu bulan Juni. Kuisioner kepuasan pelanggan terdiri dari 3 aspek yaitu Kepuasan Pelayanan, Kepuasan Produk, Kepuasan Suasana Temp. Data kepuasan pelanggan aspek pelayanan dapat dilihat pada Tabel 1.

**Tabel 1.** Data Kepuasan Pelanggan Aspek Pelayanan

| Keramahan Pelayanan       |     | Kesopanan Pelayanan       |     | Kecepatan Pelayanan      |     | Kebersihan Pelayanan     |     | Kepuasan   |     |
|---------------------------|-----|---------------------------|-----|--------------------------|-----|--------------------------|-----|------------|-----|
| Sangat ramah              | 36  | Sangat Sopan              | 20  | Cepat                    | 36  | Sangat Baik              | 17  | Puas       | 134 |
| Ramah                     | 88  | Sopan                     | 117 | Cukup Cepat              | 106 | Cukup Baik               | 125 | Tidak Puas | 71  |
| Biasa Saja                | 103 | Biasa Saja                | 110 | Biasa Saja               | 79  | Biasa Saja               | 132 | Biasa Saja | 86  |
| Kurang Ramah              | 44  | Kurang Sopan              | 44  | Kurang                   | 70  | Kurang Baik              | 17  |            |     |
|                           | 291 |                           | 291 |                          | 291 |                          | 291 |            | 291 |
| Sangat Ramah/Puas         | 44  | Sangat Sopan/Puas         | 18  | Cepat/Puas               | 28  | Sangat Baik/Puas         | 16  |            |     |
| Sangat Ramah/Tidak Puas   | 0   | Sopan/Tidak Puas          | 0   | Cepat/Tidak Puas         | 1   | Baik/Tidak Puas          | 0   |            |     |
| Sangat Ramah/Biasa Saja   | 12  | Sangat Sopan / Biasa Saja | 2   | Cepat/Biasa Saja         | 7   | Sangat Baik/ Biasa Saja  | 1   |            |     |
|                           | 56  |                           | 20  |                          | 36  |                          | 17  |            |     |
| Ramah/Puas                | 54  | Sopan /Puas               | 72  | Cukup Cepat/Puas         | 55  | Cukup Baik / Puas        | 73  |            |     |
| Ramah/Tidak Puas          | 4   | Dopan / Tidak Puas        | 8   | Cukup Cepat / Tidak Puas | 23  | Cukup Baik / Tidak Puas  | 24  |            |     |
| Ramah / Biasa Saja        | 30  | Sopan / Biasa Saja        | 37  | Cukup Cepat / Biasa Saja | 7   | Sangat Baik / Biasa Saja | 1   |            |     |
|                           | 88  |                           | 117 |                          | 106 |                          | 125 |            |     |
| Biasa Saja/ Puas          | 36  | Biasa Saja/Puas           | 42  | Biasa Saja / Puas        | 31  | Biasa Saja/Puas          | 44  |            |     |
| Biasa Saja /Tidak Puas    | 32  | Biasa Saja / Tidak Puas   | 24  | Biasa Saja / Tidak Puas  | 26  | Biasa Saja / Tidak Puas  | 38  |            |     |
| Biasa Saja/Biasa Saja     | 35  | Biasa Saja/Biasa Saja     | 44  | Biasa Saja/Biasa Saja    | 22  | Biasa Saja/Biasa Saja    | 50  |            |     |
|                           | 103 |                           | 110 |                          | 79  |                          | 132 |            |     |
| Kurang Ramah / Puas       | 0   | Kurang Sopan / Puas       | 2   | Kurang Cepat / Puas      | 20  | Kurang Baik / Puas       | 1   |            |     |
| Kurang Ramah / Tidak Puas | 35  | Kurang Sopan/Tidak Puas   | 39  | Kurang Cepat/Tiak Puas   | 21  | Kurang Baik / Tidak Puas | 9   |            |     |
| Kurang Ramah/Biasa Saja   | 9   | Kurang Sopan/Biasa Saja   | 3   | Kurang Cepat/Biasa Saja  | 29  | Kurang Baik/Biasa Saja   | 7   |            |     |
|                           | 44  |                           | 44  |                          | 70  |                          | 17  |            |     |

Dalam pembuatan pohon keputusan, yang harus dilakukan adalah menghitung jumlah kasus, jumlah kasus untuk keputusan “Puas”, jumlah kasus untuk keputusan “Biasa Saja” jumlah kasus untuk keputusan “Tidak Puas” dan kasus yang dibagi berdasarkan atribut. Setelah itu, lakukan perhitungan *Gain* untuk setiap atribut.

*Entropy* Keramahan Pelayanan = Sangat Ramah

$$\begin{aligned}
 &= \left( -\frac{0}{56} * \log_2 \left( \frac{0}{56} \right) \right) + \left( -\frac{12}{56} * \log_2 \left( \frac{12}{56} \right) \right) + \left( -\frac{44}{56} * \log_2 \left( \frac{44}{56} \right) \right) \\
 &= 0 + 0,476227 + 0,273368 \\
 &= 0,749595
 \end{aligned}$$

*Entropy* Keramahan Pelayanan = Ramah

$$\begin{aligned}
&= \left( -\frac{4}{88} * \log_2 \left( \frac{4}{88} \right) \right) + \left( -\frac{30}{88} * \log_2 \left( \frac{30}{88} \right) \right) + \left( -\frac{54}{88} * \log_2 \left( \frac{54}{88} \right) \right) \\
&= 0,202701 + 0,529275 + 0,432334 \\
&= 1,16431
\end{aligned}$$

*Entropy* Keramahan Pelayanan = Biasa Saja

$$\begin{aligned}
&= \left( -\frac{32}{103} * \log_2 \left( \frac{32}{103} \right) \right) + \left( -\frac{35}{103} * \log_2 \left( \frac{35}{103} \right) \right) + \left( -\frac{36}{103} * \log_2 \left( \frac{36}{103} \right) \right) \\
&= 0,523961 + 0,529152 + 0,530065 \\
&= 1,583178
\end{aligned}$$

*Entropy* Keramahan Pelayanan = Kurang Ramah

$$\begin{aligned}
&= \left( -\frac{35}{44} * \log_2 \left( \frac{35}{44} \right) \right) + \left( -\frac{9}{44} * \log_2 \left( \frac{9}{44} \right) \right) + \left( -\frac{0}{44} * \log_2 \left( \frac{0}{44} \right) \right) \\
&= 0,262618 + 0,488308 + 0 \\
&= 0,730926
\end{aligned}$$

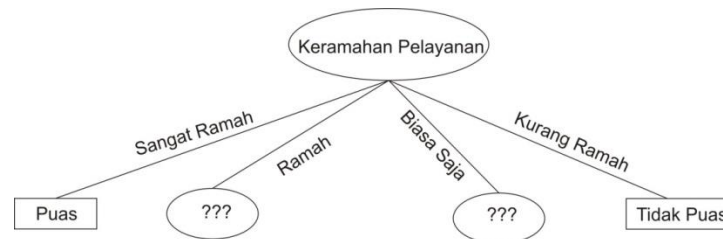
Setelah menghitung *Entropy* Atribut Keramahan Pelayanan dilanjutkan dengan menghitung *Gain* Atribut Keramahan Pelayanan dengan menggunakan persamaan yang sama untuk menghitung *Gain* hasil perhitungan sebagai berikut :

$$\begin{aligned}
&= 1,519345 - \left( \frac{56}{291} * 1,749595 \right) + \left( \frac{88}{291} * 1,164310 \right) + \left( \frac{103}{291} * 1,583178 \right) + \left( \frac{44}{291} * 0,730926 \right) \\
&= 1,519345 - 0,144252 + 0,352094 + 0,560369 + 0,110518 \\
&= 1,519345 - 1,167233 \\
&= 0,352112
\end{aligned}$$

Setelah semua perhitungan *Entropy* dan *Gain* Aspek Pelayanan di lakukan maka di hasilkan data pada Tabel 2.

**Tabel 2.** *Entropy* dan *Gain* Aspek Kepuasan Pelayanan

| Atribut              | Value        | Jml Kasus | Tidak Puas | Biasa Saja | Puas | <i>Entropy</i> | <i>Gain</i> |
|----------------------|--------------|-----------|------------|------------|------|----------------|-------------|
| Total                |              | 291       | 71         | 86         | 124  | 1,519345       |             |
| Keramahan Pelayanan  |              |           |            |            |      |                | 0,352112    |
|                      | Sangat Ramah | 56        | 0          | 12         | 44   | 0,749595       |             |
|                      | Ramah        | 88        | 4          | 30         | 54   | 1,16431        |             |
|                      | Biasa Saja   | 103       | 32         | 35         | 36   | 1,583178       |             |
|                      | Kurang Ramah | 44        | 35         | 9          | 0    | 0,730926       |             |
| Kesopanan Pelayanan  |              |           |            |            |      |                | 0,3208032   |
|                      | Sangat Sopan | 20        | 0          | 2          | 18   | 0,468996       |             |
|                      | Sopan        | 117       | 8          | 37         | 72   | 1,220925       |             |
|                      | Biasa Saja   | 110       | 24         | 44         | 42   | 1,538347       |             |
|                      | Kurang Sopan | 44        | 39         | 3          | 2    | 0,621122       |             |
| Kecepatan Pelayanan  |              |           |            |            |      |                | 0,080713    |
|                      | Cepat        | 36        | 1          | 7          | 28   | 0,884997       |             |
|                      | Cukup        | 106       | 23         | 28         | 55   | 1,476763       |             |
|                      | Cepat        | 79        | 26         | 22         | 31   | 1,570882       |             |
|                      | Biasa Saja   | 70        | 21         | 29         | 20   | 1,516359       |             |
| Kebersihan Pelayanan |              |           |            |            |      |                | 0,1136931   |
|                      | Sangat Baik  | 17        | 0          | 1          | 16   | 0,322757       |             |
|                      | Baik         | 125       | 24         | 28         | 73   | 1,393766       |             |
|                      | Cukup Baik   | 132       | 38         | 50         | 44   | 1,575993       |             |
|                      | Biasa Saja   | 17        | 9          | 7          | 1    | 1,253297       |             |
|                      | Kurang       |           |            |            |      |                |             |



**Gambar 2.** Pohon keputusan hasil perhitungan *node 1*

Adapun pohon keputusan yang terbentuk dari pencarian *Node 1* (Gambar 2) adalah Sangat Ramah memiliki keputusan “Puas”, dan Kurang Ramah memiliki keputusan “Tidak Puas”, sedangkan ramah dan Biasa Saja masih perlu dilakukan perhitungan lagi. *Entropy* dan *gain* keramahan pelayanan = ramah dapat dilihat pada Tabel 3.

**Tabel 3.** *Entropy* dan *Gain* Keramahan Pelayanan = Ramah

| Atribut                     | Value        | Jml Kasus | Tidak Puas | Biasa Saja | Puas | Entropy  | Gain      |
|-----------------------------|--------------|-----------|------------|------------|------|----------|-----------|
| Keramahan Pelayanan = Ramah |              | 88        | 4          | 30         | 54   | 1,16431  |           |
| Kesopanan Pelayanan         |              |           |            |            |      |          | 0,0765788 |
|                             | Sangat Sopan | 7         | 0          | 0          | 7    | 0        |           |
|                             | Sopan        | 40        | 2          | 14         | 24   | 1,188376 |           |
|                             | Biasa Saja   | 36        | 1          | 14         | 21   | 1,127101 |           |
|                             | Kurang Sopan | 5         | 1          | 2          | 2    | 1,521928 |           |
| Kecepatan Pelayanan         |              |           |            |            |      |          | 0,1262164 |
|                             | Cepat        | 12        | 0          | 1          | 11   | 0,413817 |           |
|                             | Cukup        | 35        | 1          | 10         | 24   | 1,036186 |           |
|                             | Cepat        | 21        | 0          | 10         | 11   | 0,998363 |           |
|                             | Biasa Saja   | 20        | 3          | 9          | 8    | 1,457717 |           |
| Kebersihan Pelayanan        |              |           |            |            |      |          | 0,210847  |
|                             | Sangat Baik  | 4         | 0          | 0          | 4    | 0        |           |
|                             | Cukup Baik   | 41        | 0          | 8          | 33   | 0,712064 |           |
|                             | Biasa Saja   | 36        | 2          | 18         | 16   | 1,25163  |           |
|                             | Kurang Baik  | 7         | 2          | 4          | 1    | 1,378784 |           |



**Gambar 3.** Pohon Keputusan Hasil Perhitungan *Node 1.1*

Adapun pohon keputusan yang terbentuk dari pencarian *Node 1.1* (Gambar 3) Kebersihan Pelayanan Sangat Baik memiliki keputusan “Puas”, dan Kebersihan Cukup Baik memiliki keputusan “Puas”, sedangkan Kebersihan Biasa Saja dan Kurang Baik masih perlu dilakukan perhitungan lagi (lihat Tabel 4).

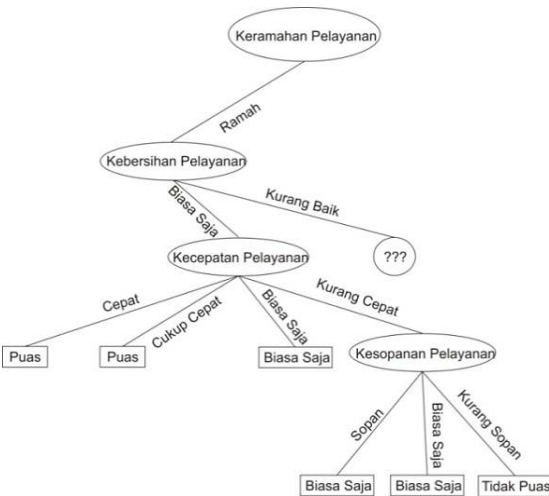
**Tabel 4. Entropy dan Gain**  
Keramahan Pelayanan = Ramah; Kebersihan Pelayanan = Biasa Saja

| Atribut                           | Value        | Jml Kasus | Tidak Puas | Biasa Saja | Puas | Entropy  | Gain     |
|-----------------------------------|--------------|-----------|------------|------------|------|----------|----------|
| Keramahan Pelayanan = Ramah       |              |           |            |            |      |          |          |
| Kebersihan Pelayanan = Biasa Saja |              | 36        | 2          | 18         | 16   | 1,25163  |          |
| Kesopanan Pelayanan               |              |           |            |            |      |          | 0,299833 |
|                                   | Sangat Sopan | 4         | 0          | 0          | 4    | 0        |          |
|                                   | Sopan        | 18        | 1          | 9          | 8    | 1,25163  |          |
|                                   | Biasa Saja   | 12        | 0          | 9          | 3    | 0,811278 |          |
|                                   | Kurang Sopan | 2         | 1          | 0          | 1    | 1        |          |
| Kecepatan Pelayanan               |              |           |            |            |      |          | 0,320392 |
|                                   | Cepat        | 3         | 0          | 0          | 3    | 0        |          |
|                                   | Cukup        | 17        | 0          | 7          | 10   | 0,977418 |          |
|                                   | Cepat        |           |            |            |      |          |          |
|                                   | Biasa Saja   | 5         | 0          | 3          | 2    | 0,97095  |          |
|                                   | Kurang Cepat | 11        | 2          | 8          | 1    | 1,095795 |          |

**Tabel 5. Entropy dan Gain**  
Keramahan Pelayanan = Ramah; Kebersihan Pelayanan = Biasa Saja;  
Kecepatan Pelayanan = Kurang Cepat

| Atribut                     | Value        | Jml Kasus | Tidak Puas | Biasa Saja | Puas | Entropy  | Gain     |
|-----------------------------|--------------|-----------|------------|------------|------|----------|----------|
| Keramahan Pelayanan = Ramah |              | 11        | 2          | 8          | 1    | 1,095795 |          |
| Kesopanan Pelayanan         |              |           |            |            |      |          | 0,413088 |
|                             | Sangat Sopan | 0         | 0          | 0          | 0    |          |          |
|                             | Sopan        | 6         | 1          | 4          | 1    |          |          |
|                             | Biasa Saja   | 4         | 0          | 4          | 0    |          |          |
|                             | Kurang Sopan | 1         | 1          | 0          | 0    |          |          |

Adapun pohon keputusan yang terbentuk dari pencarian *Node* 1.2 ada pada Gambar 4.



**Gambar 4.** Decision Tree Keramahan Pelayanan = Ramah / Kebersihan Pelayanan = Biasa Saja



Pada Tabel 5 dan Gambar 4 Kecepatan Pelayanan Cepat memiliki keputusan “Puas”, Kecepatan Pelayanan Biasa Saja memiliki keputusan “Biasa Saja”, sedangkan Kurang Cepat memiliki atribut Kesopanan Pelayanan. Kesopanan Pelayanan Sopan memiliki keputusan “Biasa Saja”, Kesopanan Pelayanan Biasa Saja memiliki keputusan “Biasa Saja”, dan Kurang Sopan memiliki keputusan “Tidak Puas”.

Aturan atau *rule* yang terbentuk dari perhitungan *Gain* dan *Entropy* adalah sebagai berikut :

1. Jika Keramahan Pelayanan = Sangat Ramah, maka pelanggan = Puas
2. Jika Keramahan Pelayanan = Kurang Ramah, maka pelanggan = Tidak Puas
3. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Sangat Baik, maka pelanggan = Puas
4. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Cukup Baik, maka pelanggan = Puas
5. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Cepat, maka pelanggan = Puas
6. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Cukup Cepat, maka pelanggan = Puas
7. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Biasa Saja, maka pelanggan = Biasa Saja
8. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Kurang Cepat dan Kesopanan Pelayanan = Sopan, maka pelanggan = Biasa Saja
9. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Kurang Cepat dan Kesopanan Pelayanan = Biasa Saja, maka pelanggan = Biasa Saja
10. Jika Keramahan Pelayanan = Ramah dan Kebersihan Pelayanan = Biasa Saja dan Kecepatan Pelayanan = Kurang Cepat dan Kesopanan Pelayanan = Kurang Sopan, maka pelanggan = Tidak Puas

Berikut data *Performance Vector* hasil pengujian model C4.5:

**Tabel 6.** Performance Vektor Hasil Pengujian C4.5

| Kelas Aktual | Puas<br>(Prediksi) | Tidak Puas<br>(Prediksi) |
|--------------|--------------------|--------------------------|
| Puas         | 20                 | 1                        |
| Tidak Puas   | 0                  | 5                        |

Dari Tabel 6, model menghasilkan akurasi 96,15%, seperti yang tercantum pada *Performance Vector*. Adapun Interpretasi *Confusion Matrix* menunjukkan kinerja model klasifikasi dengan membandingkan hasil prediksi dan data aktual. Adapun uraiannya ada pada Tabel 7.

**Tabel 7** *Confusion Matrix* Kinerja Model Klasifikasi

| Simbol              | Deskripsi                                                    | Nilai | Arti                                                   |
|---------------------|--------------------------------------------------------------|-------|--------------------------------------------------------|
| TP (True Positive)  | Data yang aktualnya “Puas” dan diprediksi “Puas”             | 20    | Model benar memprediksi pelanggan yang puas            |
| FN (False Negative) | Data yang aktualnya “Puas” tetapi diprediksi “Tidak Puas”    | 1     | Model keliru menilai pelanggan puas sebagai tidak puas |
| FP (False Positive) | Data yang aktualnya “Tidak Puas” tetapi diprediksi “Puas”    | 0     | Tidak ada kesalahan jenis ini                          |
| TN (True Negative)  | Data yang aktualnya “Tidak Puas” dan diprediksi “Tidak Puas” | 5     | Model benar mengenali pelanggan tidak puas             |

Adapun Analisis Kinerja Model C4.5, Berdasarkan metrik di atas, dapat disimpulkan pada Tabel 8. Berdasarkan pengolahan data menggunakan software RapidMiner 8.1 didapat nilai akurasi sistem sebesar 96.15%, dimana model yang telah dibentuk diuji tingkat akurasi dengan memasukkan data uji yang berasal dari data training dengan menggunakan Split Validation pada aplikasi RapidMiner 8.1 untuk menguji tingkat akurasi.

**Tabel 8.** Analisis Kerja Model C4.5

| Aspek   | Hasil  | Interpretasi                           |
|---------|--------|----------------------------------------|
| Akurasi | 96,15% | Model sangat akurat secara keseluruhan |

| Aspek                      | Hasil  | Interpretasi                                              |
|----------------------------|--------|-----------------------------------------------------------|
| <i>Precision</i><br>(Puas) | 100%   | Semua prediksi puas benar                                 |
| <i>Recall</i><br>(Puas)    | 95,20% | Model hampir tidak pernah gagal mendeteksi pelanggan puas |
| <i>Specificity</i>         | 100%   | Tidak ada kesalahan prediksi pelanggan tidak puas         |
| <i>F1-Score</i>            | 97,50% | Model sangat seimbang dan andal                           |

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan tingkat kepuasan pelanggan dengan akurasi sebesar 96,15%, dengan nilai *precision* 100% dan *recall* 95,2%. Nilai ini menunjukkan bahwa model dapat memprediksi kepuasan pelanggan dengan tingkat kesalahan yang sangat kecil. Dari hasil *confusion matrix*, hanya satu data yang salah klasifikasi, di mana pelanggan yang sebenarnya “puas” terdeteksi sebagai “tidak puas”. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang tinggi dan kesalahan klasifikasi yang minim. Kesesuaian karakteristik data dengan algoritma C4.5.

Algoritma C4.5 sangat efektif untuk *dataset* yang mengandung atribut kategorikal dan numerik sekaligus, seperti data survei kepuasan pelanggan yang memuat variabel seperti “kualitas pelayanan”, “harga”, “kecepatan layanan”, dan “respons petugas”. Dengan demikian, capaian akurasi 96,15% dalam penelitian ini lebih tinggi dari rata-rata penelitian sebelumnya, menandakan bahwa data yang digunakan memiliki kualitas yang baik (minim *noise* atau *missing value*). Pemilihan atribut relevan dan *preprocessing* yang tepat meningkatkan kemampuan model untuk membedakan kelas kepuasan. Model C4.5 bekerja sangat baik untuk konteks survei kepuasan pelanggan di mana hubungan antar variabel bersifat non-linear dan tidak selalu dapat dijelaskan secara statistik konvensional.

REFERENCES

[1] D. Telaumbanua and I. Kurniawati, “Penerapan Algoritma C4.5 Untuk Klasifikasi Kepuasan Pelanggan Pada Jasa Layanan Pengiriman,” *JoMMiT J. Multi Media dan IT*, vol. 6, no. 1, 2022.

[2] D. Yunita and I. H. Ikasari, “Perbandingan Metode Klasifikasi C4.5 dan Naïve Bayes untuk Mengukur Kepuasan Pelanggan,” *J. Inform. Univ. Pamulang*, vol. 6, no. 3, pp. 2622–4615, 2021.

[3] D. Bayu Febriyanto, L. Handoko, H. Aisyah, and Rumini, “Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop,” *J. Ris. Komput.*, vol. 5, no. 6, pp. 569–575, 2018.

[4] A. R. Damanik, I. Gunawan, R. K. Sormin, H. S. Tambunan, S. Sumarno, and S. Darma, “Analisis Metode C4.5 Dalam Mengukur Akurasi Kepuasan Mahasiswa/i Terhadap Layanan Perpustakaan,” *Digit. Transform. Technol.*, vol. 4, no. 1, pp. 699–704, 2024.

[5] C. R. Aditya Nugroho and T. Kristiana, “Penerapan Algoritma C4.5 Untuk Kepuasan Pelanggan Toko Online Parfume Chantik,” *J. Algoritm.*, vol. 3, no. 1, pp. 10–21, 2022.

[6] S. Arini Damanik and M. Safii, “Seminar Nasional Informatika (SENATIKA) Prosiding SENATIKA 2023 Implementasi Algoritma C4.5 Untuk Mengetahui Faktor Kepuasan Konsumen Pada PT Mandala Multifinance Cabang Pematang Siantar,” 2023.

[7] A. Fauzia, N. Ndruru, and A. Sindar, “Katera : Jurnal Sains dan Teknologi Penerapan Data Mining Klasifikasi Kepuasan Pelanggan Transportasi Online Menggunakan Algoritma C4.5,” vol. 1, pp. 23–30, 2024.

[8] Y. Bastian *et al.*, “Analisis Penerapan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pelanggan Indihome Pada Kota Pematangsiantar,” *Januari*, vol. 2, no. 1, pp. 62–69, 2021.

[9] S. M. Putri and S. A. Arnomo, “Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus: Hinet Batam),” *J. Inf. Syst. Res.*, vol. 1, no. 2, pp. 70–76, 2020.

[10] Y. I. Lestari, S. Defit, and Y. Yuhandri, “Prediksi Tingkat Kepuasan Pelayanan Online Menggunakan Metode Algoritma C.45,” *J. Inform. Ekon. Bisnis*, vol. 3, pp. 148–154, 2021.

[11] F. R. Wiratama, “Implementasi Algoritma C4 . 5 untuk Analisa Performa Pelayanan,” *Explor. Inform.*, vol. 6, pp. 127–135, 2017.

[12] S. G. Ahmad, D. Arifianto, and W. Suharso, “Jurnal Smart Teknologi Jurnal Smart Teknologi,” *Univ. Muhammadiyah Jember*, vol. 3, no. 5, pp. 502–510, 2022.

[13] B. Sudrajat, “Penggunaan Algoritma C4.5 Untuk Menentukan Kepuasan Pelanggan Pada Warnet Game Victory,” *J. Ilmu Tek. dan Komput.*, vol. 6, no. 1, p. 27, 2022.

[14] F. Fitrayana and K. Rizal, “Bianglala Informatika : Jurnal Komputer dan Informatika Akademi Bina Sarana Informatika Yogyakarta Penerapan Algoritma C4.5 Untuk Menentukan Kepuasan Pelanggan

- Dalam Kualitas Pelayanan Pengiriman Barang (Studi Kasus : Pt Nusantara Card Semesta Cabang Kema,” vol. 12, no. 1, p. 2024, 2024.
- [15] T. H. Toto.Hermanto, “Customer Satisfaction Analysis Using the C4.5 Algorithm at Askara Minimarket,” *Sci. J. Inf. Syst.*, vol. 2, no. 1, pp. 16–19, 2024.
- [16] A. Steffany, K. A. Aldzikri, M. A. R. Tricahya, and M. I. Arfiandono, “Penerapan Algoritma C4.5 Untuk Menganalisis Kepuasan E-Commerce Shopee,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 6, no. 6, pp. 700–706, 2023.
- [17] B. Apriliyawan and S. Lestari, “Performance Analysis Of C4.5 And Naïve Bayes Algorithm On Customer Relationship Management (CRM) In Jatimas Furniture Ltd,” *Icith*, no. December, pp. 20–27, 2020.