

## Analisis Faktor Kepuasan Konsumen Terhadap Produk Roti Pinkan Bakery & Cake dengan Algoritma C4.5

**Novita Indriyani**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Korespondensi penulis: [12noviindri@gmail.com](mailto:12noviindri@gmail.com)

**Heru Satria Tambunan**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

**Zulia Almaida Siregar**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

**Abstract.** *This study aims to determine the level of consumer satisfaction with the products of Pinkan Bakery & Cake. By knowing the level of customer satisfaction, the store can improve and increase the customer satisfaction factor if there are weaknesses and shortcomings. The method used in this research is data mining, namely the C4.5 algorithm, the source of the data is taken using a questionnaire/questionnaire at the survey stage of filling out sheets of paper to shop visitors or the surrounding community in the Bangun Village environment. The variables used are (1) Price, (2) Taste, (3) Product Quality, (4) Place Cleanliness. To facilitate the process of making a decision tree, the research test process uses Rapidminer 5.3 software. Based on all the results of the research stages that have been carried out on the application of the C4.5 Algorithm in the analysis of the Consumer Satisfaction factor for the Pinkan Bakery and Cake Products, it can be concluded that the problem of determining the consumer satisfaction factor for Bread products can be solved using data mining techniques, namely the C4.5 Algorithm. . Produces 5 rules and the accuracy rate generated by this method is 93%. From the calculation with the C4.5 algorithm, the most dominant factor is Rasa (C2) with a gain value of 0,333610402.*

**Keywords:** *Consumer Satisfaction, Data Mining, C4.5 Algorithm, Bread, Bakery & Cake.*

**Abstrak.** Penelitian ini bertujuan untuk mengetahui besar tingkat kepuasan konsumen terhadap produk roti Pinkan Bakery & Cake. Dengan mengetahui tingkat kepuasan konsumen, maka pihak toko dapat memperbaiki dan meningkatkan faktor kepuasan konsumen apabila ada kelemahan dan kekurangan. Metode yang digunakan dalam penelitian ini dengan data mining yaitu Algoritma C4.5, sumber data diambil dengan menggunakan angket/kuesioner pada tahap survei pengisian lembar kertas kepada pengunjung toko atau masyarakat sekitar di lingkungan Desa Bangun. Adapun variabel yang digunakan yaitu (1) Harga, (2) Rasa, (3) Kualitas Produk, (4) Kebersihan Tempat. Untuk memudahkan proses pembuatan pohon keputusan proses uji penelitian menggunakan software Rapidminer 5.3. Berdasarkan seluruh hasil tahapan penelitian yang telah dilakukan pada Penerapan Algoritma C4.5 pada analisis faktor Kepuasan Konsumen terhadap Produk Roti Pinkan Bakery and Cake dapat disimpulkan bahwa permasalahan menentukan faktor kepuasan konsumen terhadap produk Roti dapat diselesaikan menggunakan teknik data mining, yaitu dengan Algoritma C4.5. Menghasilkan 5 rules dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah

93%. Dari perhitungan dengan Algoritma C4.5 maka didapatkan faktor yang paling dominan adalah Rasa (C2) dengan nilai gain sebesar 0,333610402.

**Kata kunci:** Kepuasan Konsumen, Data Mining, Algoritma C4.5, Roti, Bakery & Cake.

## LATAR BELAKANG

Kepuasan konsumen menjadi hal yang utama bagi para pelaku industri pangan. Kepuasan konsumen adalah situasi yang ditunjukkan oleh konsumen ketika mereka menyadari bahwa kebutuhan dan keinginannya sesuai atau tidak dengan apa yang diharapkan. Ketika tingkat kepuasan tinggi maka ada kemungkinan bahwa mereka akan membeli kembali produk sehingga pendapatan pelaku industri tetap stabil atau bahkan meningkat menurut (Hermansyah & Roessali, 2021). Kepuasan konsumen merupakan faktor utama untuk pelaku industri dalam meningkatkan kualitas produk yang lebih baik lagi.

Pinkan Bakery & Cake merupakan toko roti yang mempunyai tujuan untuk memberikan kepuasan kepada konsumennya, antara lain dengan cara melihat dan mengetahui keinginan dan kebutuhan konsumennya. Untuk memahami perilaku para konsumennya agar tujuan yang sudah ada dapat tercapai. Usaha untuk mempertahankan konsumen tentu bukanlah hal yang mudah, apalagi hal yang sekarang sedang dihadapi oleh toko Pinkan Bakery & Cake yaitu sedang menurunnya jumlah konsumen dan pelanggan roti Pinkan Bakery & Cake. Hal tersebut menjadi sebuah ancaman yang sedang dihadapi oleh toko tersebut karena akan berdampak pada omset pendapatan.

Beberapa faktor yang mempengaruhi kepuasan konsumen terhadap produk yaitu : harga, rasa, lingkungan fisik (kebersihan tempat), ukuran produk. Dengan demikian, pihak toko harus membuat kebijakan untuk meningkatkan jumlah konsumen kembali dan kemudian mencari faktor yang paling dominan dalam mempengaruhi kepuasan konsumen.

Dalam mengetahui faktor yang paling dominan terhadap kepuasan konsumen dapat menggunakan hasil klasifikasi. "Klasifikasi adalah proses pencarian model yang dapat membedakan kelas data dengan tujuan agar model tersebut dapat digunakan untuk memprediksi kelas dari suatu obyek yang belum diketahui kelasnya. Salah satu Algoritma klasifikasi yang bisa digunakan adalah C4.5 " (Nasrullah, 2018)

Algoritma C4.5 Data Mining digunakan untuk membangun sebuah pohon keputusan yang fleksibel, menarik, dan mudah dimengerti karena dapat divisualisasikan dalam bentuk gambar. Pohon keputusan adalah salah satu metode atau cara klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Konsep dari pohon keputusan dan aturan-aturan keputusan (Ester et al., 2019).

Penelitian yang menjadi referensi dalam penelitian ini yaitu penelitian yang dilakukan oleh (Hermansyah & Roessali, 2021) membahas tentang menganalisis faktor yang mempengaruhi kepuasan konsumen menggunakan metode *Regresi Logistik*. Dari hasil perhitungan konsumen puas terhadap pembelian sawi organik di Farmers Market yang ditunjukkan oleh nilai CSI sebesar 79.52%. Dari penelitian ini sama-sama menganalisis kepuasan konsumen. Akan tetapi, perbedaan dari penelitian sebelumnya adalah menggunakan metode *Regresi Logistik* dan 4 atribut yaitu produk, harga, promosi, kepuasan pelayanan. Dari perhitungan dengan metode tersebut didapat nilai CSI sebesar 79.52% kepuasan konsumen dipengaruhi oleh produk, harga dan kualitas pelayanan

sedangkan promosi tidak berpengaruh terhadap kepuasan konsumen. Berdasarkan latar belakang tersebut, maka dilakukan lah penelitian ini.

## **KAJIAN TEORITIS**

### **1. Data Mining**

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Irnanda, 2020). Data mining juga dikenal dengan istilah *pattern recognition*, merupakan suatu metode yang digunakan untuk pengolahan data guna menemukan pola yang tersembunyi dari data yang diolah (Oktavia, Hardinata, & Irawan, 2020). Data mining merupakan metode pengolahan data berskala besar, oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum, kajian data mining membahas metode-metode seperti, asosiasi, *clustering*, klasifikasi, regresi, seleksi variable dan market basket analisis (Astuti, 2018). Banyak sekali penelitian-penelitian yang telah dilakukan dengan menggunakan data mining (Arminarahmah, GS, Bhawika, Dewi, & Wanto, 2021; Febriyati, GS, & Wanto, 2020; Gultom, Wanto, Gunawan, Lubis, & Kirana, 2021; Hanafiah & Wanto, 2020; Hutagalung et al., 2021; Sinaga, Wanto, Gunawan, Sumarno, & Nasution, 2021; Wanto et al., 2020).

### **2. Klasifikasi**

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri bisa berupa aturan “jika maka”, *decision tree* ataupun formula matematis (Febriani & Sulistiani, 2021).

Ada dua jenis model klasifikasi, yaitu (Sadewo, Windarto, Damanik, & Hartama, 2018):

- a) Pemodelan Deskriptif (*descriptive modelling*), digunakan untuk membedakan objek-objek dalam kelas-kelas yang berbeda.
- b) Pemodelan Prediktif (*predictive modelling*), digunakan untuk memprediksi label kelas *record* yang tidak diketahui.

### **3. Algoritma C4.5**

Algoritma ini adalah penerus dari algoritma sebelumnya, yaitu algoritma ID3 yang digunakan untuk menghasilkan sebuah rule berupa pohon keputusan. Algoritma ini dapat digunakan untuk klasifikasi data dan disebut sebagai *classifier statistic*. Algoritma C4.5 merupakan algoritma klasifikasi atau pengelompokan yang bersifat prediktif. Pernyataan klasifikasi pada pohon keputusan (*decision tree*) terdapat pada cabang-cabangnya dan kelas-kelas atau segmen-segmennya terdapat pada daun-daunnya. Algoritma C4.5 adalah salah satu keluarga pohon keputusan yang dapat menghasilkan sebuah pohon keputusan dan aturan atau rule untuk tujuan meningkatkan prediksi akurasi. Selain itu, model C4.5 mudah digunakan dan dipahami sebagai aturan yang diturunkan dan memiliki teknik interpretasi yang sangat mudah (Nurellisa & Fitriyah, 2020)

Secara umum, alur proses algoritma C4.5 untuk membangun pohon keputusan dalam data mining adalah sebagai berikut (Sadewo et al., 2018) :

- a) Pilih atribut sebagai simpul akar.
- b) Buat cabang untuk tiap nilai
- c) Bagi kasus dalam cabang
- d) Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki

kelas yang sama. Pemilihan atribut sebagai simpul, baik simpul akar (*root*) atau simpul internal didasarkan pada nilai *gain* tertinggi dari atribut yang ada.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, yaitu (Asmaul Husnah Nasrullah, 2018)

- a) Menyiapkan data *training*. Data *training* biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas – kelas tertentu.
- b) Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy* menggunakan rumus 2.1 :

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

Di mana :

$S$  : himpunan kasus

$n$  : jumlah partisi  $S$

$p_i$  : proporsi  $S_i$  terhadap  $S$

Setelah menghitung *entropy* dari setiap kasus selanjutnya menghitung nilai *gain* untuk pemisah objek dengan rumus 2.2.

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

Di mana :

$S$  : himpunan kasus

$A$  : Atribut

$N$  : Jumlah partisi atribut  $A$

$|S|$  : Jumlah kasus pada partisi ke-1

$|S_i|$  : Jumlah kasus dalam  $S_i$

- c) Ulangi langkah ke-2 hingga semua record terpartisi.
- d) Proses partisi pohon keputusan akan berhenti saat:
  1. Semua record dalam simpul  $N$  mendapat kelas yang sama.
  2. Tidak ada atribut di dalam record yang dipartisi lagi.
  3. Tidak ada record di dalam cabang yang kosong.

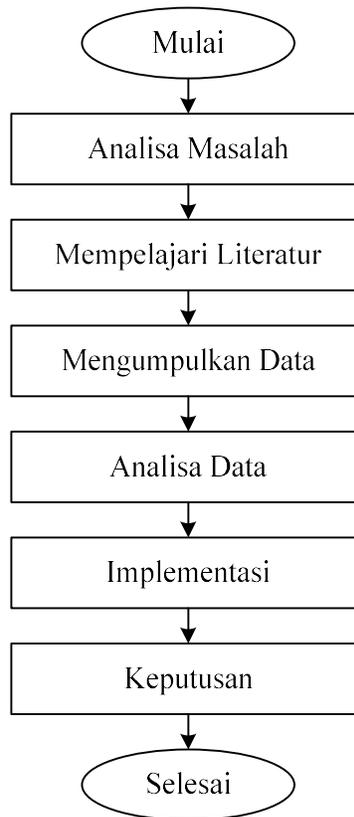
## METODE PENELITIAN

### 1. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di Toko Roti Pinkan Bakery & Cake yang berlokasi di Jl.Asahan KM.17 Desa Bangun. Waktu pengumpulan data diambil pada 1 Januari sampai akhir Juni 2022. Data Dikumpulkan dengan memberikan pertanyaan dan pernyataan melalui kuisisioner yang telah dicetak yang diberikan toko kepada konsumen yang telah berbelanja di Toko Roti Pinkan Bakery & Cake.

## 2. Rancangan Penelitian

Uraian rancangan penelitian akan disajikan dalam bentuk *flowchart* sebagai berikut.



**Gambar 1. Flowchart Rancangan Penelitian**

Berdasarkan rancangan penelitian pada gambar 1, maka masing-masing langkah dapat diuraikan sebagai berikut :

- a) **Analisa Masalah**  
Proses menganalisa suatu masalah yang terkait dengan menentukan kualitas produk terhadap kepuasan konsumen Toko dan menentukan parameter atau kriteria apa yang digunakan. Pada penelitian ini parameter atau kriteria yang digunakan yaitu Harga, Rasa, Kebersihan Lingkungan, dan Ukuran Produk.
- b) **Mempelajari Literatur**  
Penelitian ini harus didasari rujukan yang digunakan untuk mendapat informasi dan teori yang mendukung dalam penelitian.
- c) **Mengumpulkan Data**  
Data dikumpulkan dengan menggunakan kuisioner yang diisi konsumen Toko Roti Pinkan Bakery & Cake.
- d) **Analisa**  
Melakukan pengolahan data dengan menggunakan *Algoritma C4.5*. Dan data diolah menggunakan aplikasi *Ms.Excel 2010*.
- e) **Implementasi**  
Pengujian data dilakukan dengan aplikasi *Rapidminer versi 5.3*.

f) Hasil

Kesimpulan yang didapat dalam menentukan kualitas produk terhadap kepuasan konsumen Toko Pinkan Bakery & Cake yaitu diketahui aspek dominan yang membuat konsumen Toko merasa puas terhadap produk yang diberikan oleh pihak Toko.

### 3. Prosedur Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini terdapat beberapa metode sebagai berikut :

a) Penelitian Kepustakaan (*Library Research*)

Penelitian ini memanfaatkan perpustakaan untuk mencari buku-buku yang berkaitan dengan penelitian ini sebagai referensi untuk pengumpulan data.

b) Penelitian Lapangan (*Field Work Research*)

Penelitian yang dilakukan secara langsung di lapangan untuk mendapatkan data yang diinginkan dengan beberapa prosedur yaitu;

1. Wawancara

Yaitu proses pengumpulan data atau informasi melalui tatap muka antara pihak penanya dengan pihak yang ditanya atau penjawab.

2. Kuesioner

Kuisisioner adalah angket atau daftar pertanyaan yang tertulis secara tertutup atau responden hanya dapat memberikan jawaban dan secara terbuka atau responden dapat menuliskan pendapatnya berdasarkan kebutuhan data yang ingin ditanyakan. Untuk pengambilan jumlah sampel minimal. Setelah itu, peneliti membagikan kuisisioner pada pelanggan atau konsumen yang membeli Roti Pinkan Bakery & Cake.

### 4. Analisis Data

Pada penelitian ini, data diperoleh dari hasil kuesioner yang diberikan atau diisi oleh konsumen toko. Data responden dikumpulkan melalui pembagian kuesioner melalui berkas survei yang telah dicetak untuk toko. Jumlah responden yang digunakan adalah 100 responden. Variabel independen yang digunakan adalah Harga, Rasa, Lingkungan (kebersihan tempat) dan ukuran produk. Sedangkan variabel responnya adalah kepuasan yaitu SP (Sangat Puas), P (Puas), CP (Cukup puas), TP (Tidak puas). Setiap variabel diajukan pertanyaan yang berhubungan dan terkait dengan variabel. Data menggunakan jenis metode sampling dengan kuesioner yang diberikan menggunakan linker 4 yang terdiri dari SP (Sangat Puas) bernilai 4 , P (Puas) bernilai 3, CP (Cukup puas) bernilai 2 , TP (Tidak puas) bernilai 1.

**Tabel 1. Hasil Rekapitulasi Kuesioner**

No	Nama	C1 (Harga)	C2 (Rasa)	C3 (Ukuran Produk)	C4 (Kebersihan Lingkungan)	Hasil
1	R1	Sangat puas	Sangat puas	Sangat Puas	Sangat Puas	Puas
2	R2	Sangat puas	Cukup puas	Puas	Puas	Puas
3	R3	Sangat puas	Sangat puas	Sangat Puas	Sangat Puas	Puas
4	R4	Puas	Cukup puas	Cukup puas	Cukup puas	Tidak Puas
5	R5	Cukup Puas	Puas	Cukup puas	Puas	Puas
6	R6	Puas	Puas	Puas	Puas	Puas
7	R7	Sangat puas	Cukup puas	Cukup puas	Puas	Puas
8	R8	Puas	Cukup puas	Puas	Puas	Puas
9	R9	Cukup Puas	Sangat puas	Puas	Cukup puas	Puas
10	R10	Puas	Cukup puas	Cukup puas	Puas	Puas
...	...	...	...	...	...	...
100	R100	Puas	Puas	Puas	Puas	Puas

**HASIL DAN PEMBAHASAN**

**1. Proses Perhitungan Algoritma C4.5**

Pada perhitungan Algoritma C4.5 ada cara untuk memperoleh aturan pohon keputusan dan model yang dapat dilakukan dengan langkah-langkah berikut ini:

Langkah 1 : Dihitung jumlah kasus untuk bentuk keputusan Puas dan jumlah kasus untuk Tidak Puas.

Langkah 2 : Kemudian Hitung jumlah Entropy dari keseluruhan kasus yang dibagi berdasarkan kelas atribut dengan persamaan (1). Selanjutnya dilakukan perhitungan mencari nilai *Gain* untuk setiap atribut dengan persamaan (2).

Berikut ini adalah hasil perhitungan nilai entropy dan gain Node 1 diuraikan dalam tabel 2 berikut ini :

**Tabel 2. Hasil Perhitungan Node 1**

Node 1	Jumlah Responden	Puas	Tidak Puas	Entropy	Gain
Total	100	77	23	0,778011304	
<b>Harga</b>					<b>0,125077906</b>
	Sangat Puas	8	8	0	0
	Puas	81	66	15	0,691289869
	Cukup Puas	11	3	8	0,845350937
	Tidak Puas	0	0	0	0
<b>Rasa</b>					<b>0,333610402</b>
	Sangat Puas	8	8	0	0
	Puas	34	34	0	0
	Cukup Puas	38	30	8	0,74248757
	Tidak Puas	20	5	15	0,811278124

Node 1	Jumlah Responden	Puas	Tidak Puas	Entropy	Gain
Ukuran Produk					0,25866701
	Sangat Puas	9	9	0	0
	Puas	46	45	1	0,151096971
	Cukup Puas	45	23	22	0,99964375
	Tidak Puas	0	0	0	0
Kebersihan Lingkungan					0,150560056
	Sangat Puas	15	15	0	0
	Puas	72	58	14	0,710676854
	Cukup Puas	13	4	9	0,89049164
	Tidak Puas	0	0	0	0

Langkah 3 : Dari hasil perhitungan pada tabel 4.1 diperoleh nilai atribut tertinggi adalah Rasa dengan *Gain* sebesar **0,333610402**. Maka atribut Rasa dipilih sebagai *node* akar. Nilai kelas atribut sangat puas dan puas mengklasifikasikan kasus menjadi satu keputusan yaitu puas. Untuk nilai kelas atribut cukup puas dan tidak puas belum diperoleh hasil antara puas dan tidak puas, maka perlu dilakukan perhitungan lebih lanjut. Berikut ini hasil perhitungan dari kelas atribut Rasa = Tidak Puas yang ditunjukkan pada Tabel 4.2 berikut ini:

**Tabel 4.2 Hasil Perhitungan Node 1.1**

Node1.1	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Tidak Puas	20	5	15	0,811278124	
Harga					1,5
	Sangat Puas	4	0	0	0
	Puas	15	5	10	0,918295834
	Cukup Puas	5	0	5	0
	Tidak Puas	0	0	0	0
Ukuran Poduk					0,365776599
	Sangat Puas	0	0	0	0
	Puas	5	4	1	0,721928095
	Cukup Puas	15	1	14	0,353359335
	Tidak Puas	0	0	0	0
Kebersihan Lingkungan					0,183797568
	Sangat Puas	1	1	0	0
	Puas	15	4	11	0,836640742

Node1.1		Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
	Cukup Puas	4	0	4	0	
	Tidak Puas	0	0	0	0	

Dari hasil perhitungan pada Tabel 4.2 diatas, atribut yang menjadi node cabang dari Rasa-Tidak puas adalah Harga dengan nilai gain tertinggi sebesar 1,5. Untuk nilai atribut pada Rasa masih perlu dilakukan perhitungan lebih lanjut. Berikut ini hasil perhitungan Rasa-Tidak Puas dan Harga – Puas.yang ditunjukkan pada Tabel 4.3 berikut ini:

**Tabel 4.3 Hasil Perhitungan Node 1.1.1**

Node1.1.1		Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Tidak Puas, Harga-Puas		15	5	10	0,918295834	
						<b>0,595998044</b>
Ukuran Produk						
	Sangat Puas	0	0	0	0	
	Puas	4	4	0	0	
	Cukup Puas	11	1	10	0,439496987	
	Tidak Puas	0	0	0	0	
Kebersihan Lingkungan						0,146536413
	Sangat Puas	1	1	0	0	
	Puas	13	4	9	0,89049164	
	Cukup Puas	1	0	1	0	
	Tidak Puas	0	0	0	0	

Dari hasil perhitungan pada Tabel 4.3 diatas yang menjadi *node* cabang dari atribut Rasa - Tidak Puas dan Harga - Puas adalah Ukuran Produk dengan nilai *gain* tertinggi yaitu sebesar **0,595998044** dengan kelas atribut cukup puas. Kelas atribut sangat puas dan tidak puas diperoleh hasil keputusan kosong. Berikut ini adalah hasil perhitungan Rasa-Tidak puas, Harga-Puas, dan Ukuran produk-cukup puas untuk *node* 1.1.1 ditunjukkan pada tabel 4.4 berikut ini:

**Tabel 4.4 Hasil Perhitungan Node 1.1.1.1**

Node1.1.1.1		Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Tidak Puas, Harga-Puas, Ukuran produk-cukup puas		11	1	10	0,439496987	
Kebersihan Lingkungan						<b>0,439496987</b>
	Sangat Puas	1	0	0	0	
	Puas	1	1	0	0	
	Cukup Puas	0	0	0	0	

Node1.1.1.1		Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
	Tidak Puas	10	0	0	0	

Dari hasil pearahitungan Tabel 4.4 diatas, atribut yang menjadi *node* cabang dari Rasa - Tidak Puas dan Harga- Puas dan Ukuran Produk-Cukup puas adalah kebersihan lingkungan dengan *gain* sebesar **0,439496987**. Untuk nilai kelas atribut kebersihan lingkungan tidak dilakukan perhitungan lagi karena kosong. Dengan demikian *Node* 1.1.1.1 adalah node perhitungan terakhir dari atribut Rasa cabang Tidak puas. Kemudian dilakukan perhitungan kembali *Node* 1.2 atribut Rasa - Cukup Puas. Berikut Tabel 4.5 *Node* 1.2 berikut ini:

**Tabel 4.5 Hasil Perhitungan Node 1.2**

Node1.2		Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Cukup Puas		38	30	8	0,74248757	
Harga						1,262461582
	Sangat Puas	4	4	0	0	
	Puas	31	26	5	0,637387499	
	Cukup Puas	3	0	3	0	
	Tidak Puas	0	0	0	0	
Rasa - Cukup Puas		38	30	8	0,74248757	
Ukuran Poduk						0,133203816
	Sangat Puas	0	0	0	0	
	Puas	12	12	0	0	
	Cukup Puas	26	18	8	0,89049164	
	Tidak Puas	0	0	0	0	
Kebersihan Lingkungan						<b>0,37222789</b>
	Sangat Puas	3	3	0	0	
	Puas	30	27	3	0,468995594	
	Cukup Puas	5	0	5	0	
	Tidak Puas	0	0	0	0	

Dari hasil perhitungan pada tabel 4.5 diperoleh atribut yang menjadi node cabang dari Rasa-Cukup puas adalah Kebersihan Lingkungan dengan nilai gain tertinggi yaitu **0,37222789**. Untuk nilai kelas atribut dari kebesihan lingkungan dari sangat puas dan Tidak puas dan cukup puas bernilai 0 maka tidak perlu dilakukan pehitungan. Maka di atribut nilai Kebersihan lingkungan-puas masih perlu dilakukan, Berikut Tabel 4.5 *Node* 1.2.1 adalah:

**Tabel 4.6 Hasil Perhitungan Node 1.2.1**

Node1.2.1	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Cukup Puas, Harga - Puas	31	26	5	0,637387499	
Ukuran Produk					0,113982258
Sangat Puas	0	0	0	0	
Puas	11	11	0	0	
Cukup Puas	20	15	5	0,811278124	
Tidak Puas	0	0	0	0	
Kebersihan Lingkungan					<b>0,441990573</b>
Sangat Puas	2	2	0	0	
Puas	25	24	1	0,242292189	
Cukup Puas	4	0	4	0	
Tidak Puas	0	0	0	0	

Dari Tabel 4.6 Hasil Perhitungan Node 1.2.1 hasil node akar dari Rasa – Cukup Puas dan Harga- Puas adalah Kebersihan Lingkungan dan didapat hasil perhitungan tertinggi **0,441990573**. Dari nilai keals atribut Tidak puas bernilai kosong yang artinya tidak perlu dihitung lagi.Sedangkan nilai kelas atribut sangat puas dan cukup puas juga belum diketahui hasil akhirnya puas atau tidak puas.

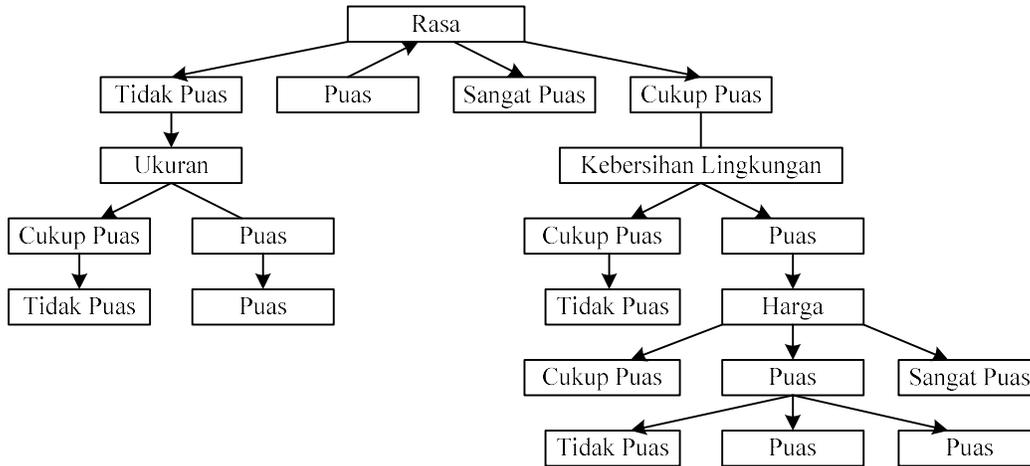
**Tabel 4.7 Hasil Perhitungan Node 1.2.1.1**

Node1.2.1.1	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
Rasa - Cukup Puas, Harga - Puas, Kebersihan Lingkungan - Puas	25	24	1	0,242292189	
Ukuran Poduk					0,026426546
Sangat Puas	0	0	0	0	
Puas	9	9	0	0	
Cukup Puas	16	15	1	0,337290067	
Tidak Puas	0	0	0	0	

Dari hasil perhitungan pada Tabel 4.7 diatas, atribut yang menjadi node cabang dari rasa-cukup puas, harga-puas, Kerbersihan lingkungan adalah Ukuran Produk dengan nilai gain 0,026426546 dengan nilai kelas atribut dari cukup puas. Untuk nilai kelas atribut sangat puas, puas, dan tidak puas bernilai kosong, maka tidak dilakukan perhitungan *node* selanjutnya. *Node* 1.2.1.1 adalah *node* cabang terakhir yang terbentuk dari C4.5.

## 2. Pohon Keputusan

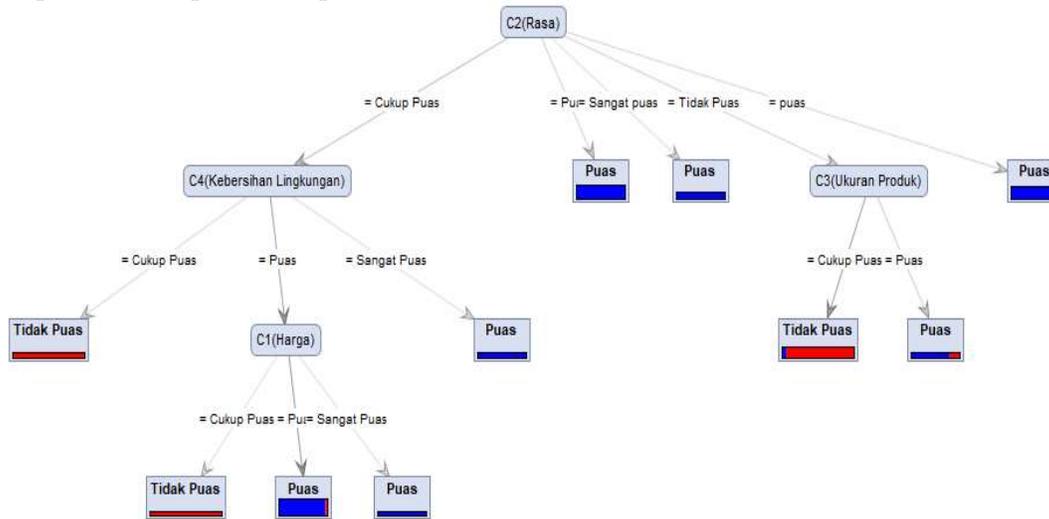
Hasil akhir dari perhitungan node-node dengan menyatukan dan menyusun gambarnya menjadi keputusan dapat dilihat pada gambar 2 berikut.



Gambar 2. Hasil Akhir Pohon Keputusan

## 3. Pengujian

Hasil pengolahan data dengan model pohon keputusan sesuai dengan *Software Rapidminer*, dapat dilihat pada Gambar 3 berikut.



Gambar 3. Pohon Keputusan dengan Rapidminer

Setelah dilakukan perhitungan dan pengujian data pada masing-masing atribut dengan Algoritma C4.5, maka akan didapatkan pola pohon keputusan akhir. Pola keputusan akhir tertinggi adalah variabel Rasa. Variabel Rasa terdiri dari 4 variabel responden yang terdiri dari 2 cabang variabel yang sudah dihasilkan puas dan yang belum diketahui hasilnya.

```

Tree

C2(Rasa) = Cukup Puas
| C4(Kebersihan Lingkungan) = Cukup Puas: Tidak Puas {Puas=0, Tidak Puas=5}
| C4(Kebersihan Lingkungan) = Puas
| | C1(Harga) = Cukup Puas: Tidak Puas {Puas=0, Tidak Puas=2}
| | C1(Harga) = Puas: Puas {Puas=25, Tidak Puas=1}
| | C1(Harga) = Sangat Puas: Puas {Puas=2, Tidak Puas=0}
| C4(Kebersihan Lingkungan) = Sangat Puas: Puas {Puas=3, Tidak Puas=0}
C2(Rasa) = Puas: Puas {Puas=19, Tidak Puas=0}
C2(Rasa) = Sangat puas: Puas {Puas=7, Tidak Puas=0}
C2(Rasa) = Tidak Puas
| C3(Ukuran Produk) = Cukup Puas: Tidak Puas {Puas=1, Tidak Puas=14}
| C3(Ukuran Produk) = Puas: Puas {Puas=4, Tidak Puas=1}
C2(Rasa) = puas: Puas {Puas=16, Tidak Puas=0}
    
```

**Gambar 4. Rules Decision Tree Pada RapidMiner**

Gambar 4 menunjukkan hasil deskripsi secara lengkap dari pohon keputusan (*decision tree*) yang telah terbentuk dengan menggunakan Algoritma C4.5. Dari hasil deskripsi juga menunjukkan bahwa penggunaan data mining algoritma C4.5 baik digunakan dalam proses menggali data (data mining proses) untuk narik beberapa kesimpulan yang divisualisasikan dengan pohon keputusan (*decision tree*). Berikut ini rules yang dihasilkan dari pohon keputusan tersebut.

**Tabel 4.8 Rule yang akan dihasilkan**

No	Rules	Keputusan
1	Jika C2=Puas dan C3=Puas dan C1= Puas	Puas
2	Jika C2= Cukup Puas dan C4= Cukup Puas dan C1= Cukup Puas	Tidak Puas
3	Jika C1= Sangat Puas dan C4= Sangat Puas dan C2= Sangat Puas	Puas
4	Jika C2 = Tidak Puas	Tidak Puas
5	Jika C3= Cukup Puas	Tidak Puas

#### 4. Tingkat Akurasi

Hasil penerapan Algoritma C4.5 menggunakan *software Rapidminer* dengan operator *X Validation* diperoleh nilai *Accuracy* yaitu sebesar 87.88% dan *classification\_error* sebesar 12% artinya bahwa *rule* yang dihasilkan tingkat kebenarannya mendekati 100%, seperti yang terlihat pada Gambar 6.

accuracy: 93.00% +/- 10.05% (mikro: 93.00%)			
	true Puas	true Tidak Puas	class precision
pred. Puas	75	5	93.75%
pred. Tidak Puas	2	18	90.00%
class recall	97.40%	78.26%	

**Gambar 5. Nilai Akurasi Algoritma C4.5**

PerformanceVector		
PerformanceVector:		
accuracy: 93.00% +/- 10.05% (mikro: 93.00%)		
ConfusionMatrix:		
True:	Puas	Tidak Puas
Puas:	75	5
Tidak Puas:	2	18

**Gambar 6. Nilai Performance Vector Algoritma C4.5**

Berdasarkan pengolahan data menggunakan *software Rapidminer* didapat nilai akurasi sistem sebesar 93% , artinya bahwa rule yang dihasilkan tingkat kebenaran mendekati 100%. Dimana model yang telah dibentuk diuji tingkat akurasinya dengan memasukan atau uji yang berasal dari data training dengan menggunakan *split validation* pada aplikasi *rapidminer 5.3* untuk menguji tingkat akurasi.

## KESIMPULAN DAN SARAN

Berdasarkan seluruh hasil tahapan penelitian yang telah dilakukan pada Penerapan Algoritma C4.5 pada analisis faktor Kepuasan Konsumen terhadap Produk Roti Pinkan Bakery and Cake dapat disimpulkan bahwa permasalahan menentukan faktor kepuasan konsumen terhadap produk Roti dapat diselesaikan menggunakan teknik data mining, yaitu dengan Algoritma C4.5. Menghasilkan 5 rules dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 93%. Dari perhitungan dengan Algoritma C4.5 maka didapatkan faktor yang paling dominan adalah Rasa (C2) dengan nilai gain sebesar 0,333610402.

## DAFTAR REFERENSI

- Arminarahmah, N., GS, A. D., Bhawika, G. W., Dewi, M. P., & Wanto, A. (2021). Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms. *IOP Conf. Series: Materials Science and Engineering*, 1071(1), 012018. <https://doi.org/10.1088/1757-899X/1071/1/012018>
- Asmaul Husnah Nasrullah. (2018). penerapan metode C45 untuk klasifikasi mahasiswa drop out. *ILKOM Jurnal Ilmiah*, 10, 244–250.
- Ester, R., Sihombing, A., Hardinata, J. T., Masruro, Z., Bangsa, S. T., & Utara, S. (2019). Penerapan Algoritma C4 . 5 Data Mining Dalam Mengukur Tingkat Kepuasan Masyarakat Kecamatan Siantar Terhadap Perbaikan Jalan 2019. 2(2), 91–98.
- Febriani, S., & Sulistiani, H. (2021). ANALISIS DATA HASIL DIAGNOSA UNTUK KLASIFIKASI GANGGUAN. 2(4), 89–95.
- Febriyati, N. A., GS, A. D., & Wanto, A. (2020). GRDP Growth Rate Clustering in Surabaya City uses the K- Means Algorithm. *International Journal of Information System & Technology*, 3(2), 276–283.
- Gultom, W. T. C., Wanto, A., Gunawan, I., Lubis, M. R., & Kirana, I. O. (2021). Application ofThe Levenberg Marquardt Method In Predict The Amount of Criminality in Pematangsiantar City. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 21–29.

<https://doi.org/10.47709/cnahpc.v3i1.926>

- Hanafiah, M. A., & Wanto, A. (2020). Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra. *International Journal of Information System & Technology*, 3(2), 315–322.
- Hermansyah, D., & Roessali, W. (2021). *Analisis Faktor yang Mempengaruhi Kepuasan Konsumen Pada Pembelian Sawi Organik di Farmers Market Semarang Analyze of Factors Affecting Consumer Satisfaction on The Purchase of Organic Mustard Green in Famers Market Semarang Kepuasan konsumen mempunyai vi.* 19(2), 177–188.
- Hutagalung, J., Ginantra, N. L. W. S. R., Bhawika, G. W., Parwita, W. G. S., Wanto, A., & Panjaitan, P. D. (2021). COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm. *Journal of Physics: Conference Series*, 1783(1), 012027. <https://doi.org/10.1088/1742-6596/1783/1/012027>
- Irnanda. (2020). Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 304–308.
- Nasrullah, A. H. (2018). *PENERAPAN METODE C4 .5 UNTUK KLASIFIKASI MAHASISWA.* 10, 244–250.
- Nurellisa, L., & Fitriana, D. (2020). Analisis Rekomendasi Calon Debitur Motor pada PT.XYZ menggunakan Algoritma C 4.5. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(4), 673. <https://doi.org/10.25126/jtiik.2020742080>
- Oktavia, R., Hardinata, J. T., & Irawan, I. (2020). Penerapan Metode Algoritma K-means Dalam Pengelompokan Angka Harapan Hidup Saat Lahir Menurut Provinsi. *Kesatria: Jurnal Penerapan ...*, 1(4), 154–161.
- Sadewo, M. G., Windarto, A. P., Damanik, I. S., & Hartama, D. (2018). Penerapan C4.5 Untuk Memprediksi Kepuasan Pasien Terhadap Sistem Informasi Manajemen Rumah Sakit. *Ready Star*, 1(1), 21–30.
- Sinaga, T. H., Wanto, A., Gunawan, I., Sumarno, S., & Nasution, Z. M. (2021). Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 9–20. <https://doi.org/10.47709/cnahpc.v3i1.923>
- Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupulu, D., ... Prianto, C. (2020). *Data Mining : Algoritma dan Implementasi.* Yayasan Kita Menulis.