



Analisis Preferensi Konsumen dalam Memilih Produk Hortikultura Menggunakan Metode Algoritma C45 dan Naive Bayes

Said Ridho*, Depi Rusda

Program Studi Sistem Informasi/Fakultas Ilmu Komputer – Universitas Darwan Ali
Sampit, Indonesia

*saidridho66@gmail.com

Abstract – Currently, agriculture and horticulture play a crucial role in the economies of many countries, including Indonesia. In horticulture, tomatoes and chili peppers are two types of vegetables that are commonly cultivated. Both of these food items hold significant importance in traditional Indonesian cuisine. The abundant availability of tomatoes and chili peppers provides consumers with more choices according to their needs. However, if the availability of tomatoes and chili peppers is limited, consumers will have restricted options in their consumption behavior. This study aims to analyze consumers' preferences for the combination of attributes related to tomatoes and chili peppers at the Keramat Baamang Sampit Market. Two algorithmic methods, C4.5 and Naive Bayes, are employed in this consumer preference analysis using the RapidMiner software for classification. The analysis results using the C4.5 algorithm to select tomatoes show an accuracy rate between 60.00% to 63.06%, while the Naive Bayes algorithm has an accuracy rate between 45.00% to 49.05%. Therefore, it can be concluded that in this case, the C4.5 algorithm demonstrates a higher accuracy rate. However, the analysis of consumer preferences for selecting chili peppers shows that the C4.5 algorithm has an accuracy rate of approximately 89.00% to 89.05%, whereas the Naive Bayes algorithm has an accuracy rate of approximately 89.05% to 90.00%. Consequently, the Naive Bayes algorithm exhibits a higher accuracy rate.

Abstrak – Saat ini, pertanian dan hortikultura memainkan peran yang sangat penting dalam perekonomian banyak negara, termasuk Indonesia. Dalam hortikultura, tomat dan cabai merupakan dua jenis tanaman sayur yang sering dibudidayakan. Kedua bahan makanan ini memegang peran penting dalam makanan tradisional Indonesia. Ketersediaan tomat dan cabai rawit yang melimpah memberikan lebih banyak pilihan bagi konsumen sesuai dengan kebutuhan mereka. Namun, jika ketersediaan tomat dan cabai terbatas, konsumen akan memiliki pilihan yang terbatas dalam perilaku konsumennya. Penelitian ini bertujuan menganalisis preferensi konsumen terhadap kombinasi atribut tomat dan cabai rawit di Pasar Keramat Baamang Sampit. Dalam analisis preferensi konsumen ini, digunakan dua metode algoritma yaitu C4.5 dan Naive Bayes dengan menggunakan software RapidMiner untuk melakukan klasifikasi. Hasil analisis menggunakan algoritma C4.5 untuk memilih tomat menunjukkan tingkat akurasi antara 60,00% hingga 63,06%, sedangkan algoritma Naive Bayes memiliki akurasi antara 45,00% hingga 49,05%. Dengan demikian, dapat disimpulkan bahwa dalam hal ini, algoritma C4.5 memiliki tingkat akurasi yang lebih tinggi. Namun, analisis preferensi konsumen untuk memilih cabai rawit menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi sekitar 89,00% hingga 89,05%, sementara algoritma Naive Bayes memiliki tingkat akurasi sekitar 89,05% hingga 90,00%. Oleh karena itu, algoritma Naive Bayes menunjukkan tingkat akurasi yang lebih tinggi.

Kata Kunci – Tomat; cabai rawit; c45; naive bayes; hortikultura.

I. PENDAHULUAN

PADA saat ini, pertanian dan hortikultura memainkan peran yang sangat penting dalam perekonomian banyak negara, termasuk di Indonesia. Hortikultura sendiri merujuk pada budidaya tanaman-tanaman hias, buah-buahan, dan sayuran yang dilakukan secara intensif dan berorientasi pada hasil produksi yang berkualitas baik [1]. Salah satu jenis tanaman sayur yang

sering dibudidayakan dalam hortikultura adalah tomat dan cabai rawit. Tomat dan cabai rawit adalah dua bahan pokok yang penting bagi setiap rumah tangga. Makanan tradisional Indonesia umumnya menggunakan tomat dan cabai sebagai bahan utama, dan seringkali tomat dan cabai dapat digunakan secara serupa dalam berbagai masakan salah satunya sambal.

Apabila tomat dan cabai tersedia dalam jumlah yang melimpah, konsumen akan memiliki lebih banyak opsi untuk memilih produk sesuai dengan kebutuhan mereka. Namun, jika tomat dan cabai terbatas atau jumlahnya terbatas, konsumen akan memiliki keterbatasan

Naskah diterima 02-05-2023, revisi 10-01-2024, terbit online 29-03-2024. Emitor merupakan Jurnal Teknik Elektro – Universitas Muhammadiyah Surakarta yang terakreditasi dengan Sinta 3 beralamat di <https://journals2.ums.ac.id/index.php/emitor/index>.

dalam memilih atau menentukan perilaku konsumen mereka. Perilaku pembelian konsumen merupakan tindakan yang dilakukan oleh konsumen saat mereka memilih di antara beberapa pilihan produk atau jasa yang tersedia untuk dibeli dan dimiliki. Perilaku konsumen melibatkan kemampuan dan usaha konsumen dalam memenuhi kebutuhan mereka. Konsumen mempertimbangkan banyak faktor sebelum membuat keputusan dalam melakukan pembelian [2].

Dalam penelitian Mhd Wendico Herdian, Riki Winanjaya, dan Susiani dengan judul "Penerapan Algoritma C4.5 Dalam Klasifikasi Kualitas Sayur Kol Di Kabupaten Simalungun". Berdasarkan penelitiannya, hasil menunjukkan bahwa penerapan Data Mining dengan menggunakan Algoritma C4.5 memungkinkan untuk diterapkan. Hasil pengujian menggunakan perangkat lunak RapidMiner menghasilkan tingkat akurasi mencapai 100% [3]. Selanjutnya berdasarkan jurnal penelitian milik Muhammad Syam Firdaus, Aji Primajaya, dan Asep Jamaludin yang berjudul "Implementasi Algoritma C4.5 untuk Prediksi Penanaman Cabai Merah" membahas penggunaan algoritma C4.5 dalam memprediksi penanaman cabai merah. Berdasarkan penelitian yang dilakukan, jurnal ini menyimpulkan bahwa algoritma C4.5 dapat digunakan untuk memprediksi penanaman cabai melalui penggunaan teknik seperti cross-validation dan confusion matrix. Dalam penelitian ini, algoritma C4.5 diimplementasikan dengan metode k-fold cross-validation, dengan nilai k sebesar 10. Hasil pengujian menunjukkan bahwa pembagian dataset memiliki pengaruh terhadap performa algoritma C4.5, dengan variasi dalam nilai akurasi, presisi, recall, dan f1-score untuk setiap skenario pengujian. Rata-rata akurasi yang diperoleh adalah 0,9944444, rata-rata presisi adalah 0,989331436, rata-rata recall adalah 1, dan rata-rata f1-score adalah 0,994557571. Selanjutnya, penelitian ini melakukan pengujian prediksi menggunakan aturan yang dihasilkan dari pohon keputusan yang telah dibuat. Hasilnya menunjukkan bahwa prediksi memiliki tingkat keakuratan yang baik, dengan sebagian besar data terprediksi dengan benar. Akurasi prediksi yang diperoleh adalah sebesar 0,978947368. Secara keseluruhan, jurnal ini menunjukkan bahwa algoritma C4.5 efektif dalam memprediksi penanaman cabai merah [4]. Berdasarkan Penelitian yang dilakukan oleh Ardian Nainggolan, Humuntal Rumapea, Arina Prima Silalahi, Lamria Sidauruk, dan Marzuki Sinambela pada tahun 2022 yang berjudul "Identifikasi Penyakit Tanaman Tomat Berdasarkan Citra Penyakit Menggunakan Metode GLCM dan Naive Bayes Classifier" membahas penggunaan metode GLCM dan Naive Bayes Classifier dalam mengidentifikasi penyakit pada tanaman tomat berdasarkan citra. Dalam jurnal ini,

penelitian ini menggunakan metode GLCM untuk melakukan ekstraksi fitur pada citra penyakit tanaman tomat, dengan fokus pada nilai energi dan entropy citra. Selanjutnya, metode Naive Bayes Classifier digunakan untuk mengklasifikasikan citra penyakit ke dalam kelas-kelas penyakit yang ada. Hasil pengujian menunjukkan bahwa pendekatan ini berhasil mencapai akurasi sebesar 80% [5].

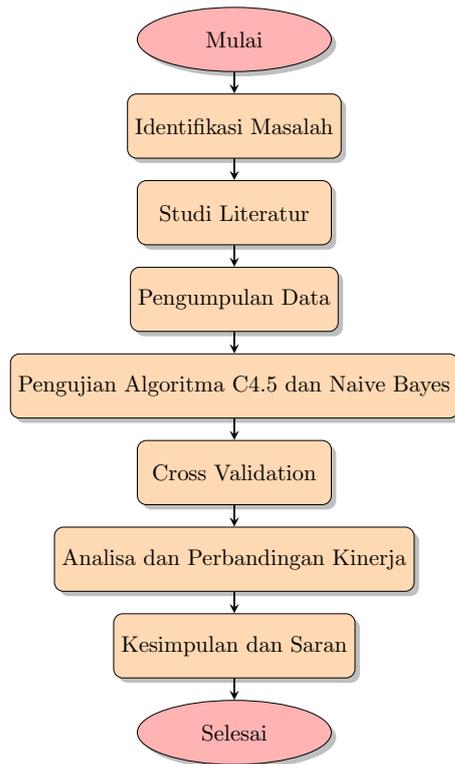
Penelitian ini bertujuan untuk mengungkap dan menganalisis preferensi konsumen terhadap tomat dan cabai rawit berdasarkan atribut-atribut tertentu seperti warna, ukuran, bentuk, tekstur dan rasa. Hal ini sangat penting dalam memenuhi kebutuhan pangan sesuai dengan keinginan konsumen. Untuk menangani permasalahan di atas, dapat dilakukan analisis preferensi konsumen terhadap kombinasi atribut tomat dan cabai rawit di Pasar Keramat Baamang Sampit dengan menggunakan metode algoritma C4.5 dan Naive Bayes. Metode algoritma C4.5 dapat digunakan untuk membangun model keputusan berdasarkan atribut-atribut yang diberikan. Algoritma ini dapat membantu dalam menentukan atribut yang paling relevan dalam memprediksi preferensi konsumen terhadap kombinasi atribut tomat dan cabai rawit. Dengan memanfaatkan algoritma C4.5, dapat dilakukan analisis pemilihan atribut yang signifikan dan membangun pohon keputusan yang memberikan pemahaman yang lebih baik tentang preferensi konsumen. Selain itu, metode Naive Bayes bisa digunakan untuk mengklasifikasikan preferensi konsumen berdasarkan kombinasi atribut yang diberikan. Metode ini bekerja dengan asumsi bahwa setiap atribut independen satu sama lain, sehingga dapat menghasilkan probabilitas klasifikasi yang akurat. Dengan penerapan metode Naive Bayes, diharapkan dapat menganalisis dan mengidentifikasi atribut yang menjadi prioritas bagi konsumen saat membeli tomat dan cabai rawit.

II. METODE PENELITIAN

i. Tahap-tahap Penelitian

Pada tahap penelitian dapat dilihat pada gambar 1. Pada gambar 1 menunjukkan tahapan penelitian yang ada 8 tahap yaitu tahap identifikasi masalah, studi literatur, pengumpulan data, pengujian algoritma c45 dan naive bayes menggunakan software rapid miner, cross validation, analisa dan perbandingan kinerja kedua algoritma, dan terakhir kesimpulan dan saran.

Pada tahap awal, yang akan dilakukan pertama kali adalah identifikasi masalah, tujuan identifikasi permasalahan yang akan diteliti adalah untuk mendapatkan pemahaman yang mendalam tentang permasalahan tersebut dan kemampuan untuk mengatasi masalah tersebut dengan efektif dan tepat. Selanjutnya tahap studi



Gambar 1: Tahapan penelitian

kiteratur pada langkah ini, yang dilakukan adalah eksplorasi terhadap berbagai landasan teori yang didapat dari sumber-sumber seperti jurnal, buku, dan referensi lainnya. Tujuannya adalah untuk menyempurnakan penelitian ini dari segi konsep dan teori, sehingga penelitian ini memiliki pedoman yang baik dan relevan. Tahapan berikutnya adalah pengumpulan data. Dalam proses pengumpulan data, menggunakan kuesioner secara online yang disediakan melalui Google Form. Kuesioner tersebut diberikan kepada 90 responden yang merupakan konsumen yang pernah membeli tomat dan cabai rawit di Pasar Keramat Baamang Sampit dan pedagang berjumlah 10 orang. Jadi jumlah data pada penelitian ini berjumlah 100. Selanjutnya dari data kuesioner akan diolah, dalam tahapan ini, data yang telah dikumpulkan akan melalui proses penyortiran terlebih dahulu. Tujuannya adalah untuk memilih dan mengambil hanya data yang diperlukan untuk analisis lebih lanjut. Setelah itu data akan di uji menggunakan kedua metode algoritma c45 dan naive bayes menggunakan software rapid miner. Tujuan dari pengujian adalah untuk mengevaluasi kinerja klasifikasi pada kelas yang telah ditentukan [6]. Dalam penelitian ini, pengujian dilakukan menggunakan metode K-Fold Validation untuk membagi data menjadi subset yang akan digunakan dalam pengujian. Untuk proses tersebut, operator Cross Validation digunakan dengan variasi nilai K-Fold Validation yang berbeda. K-Fold Validation adalah metode yang digunakan untuk membagi data menjadi beberapa

subset atau "fold" [7]. Nilai K menentukan jumlah subset yang akan dibuat. Setiap subset akan digunakan secara bergantian sebagai data pengujian, sementara subset lainnya digunakan sebagai data pelatihan [8]. Selanjutnya dari hasil pengujian akan dilakukan perbandingan dari kinerja kedua algoritma, tahap terakhir mengambil kesimpulan berdasarkan hasil yang di peroleh.

ii. Algoritma c45

Algoritma C4.5 adalah algoritma yang mengembangkan *Decision Tree* dimana mudah dimengerti dan lebih efisien dalam menginduksi data [9]. Konsep pemikiran dari *Decision Tree* dimulai dari atribut dengan nilai *gain* tertinggi yang disebut sebagai *root*, kemudian dilanjutkan ke cabang-cabang di bawahnya, hingga mencapai atribut dengan nilai *gain* terendah yang disebut sebagai *daun* [10]. Dalam algoritma ini, *node* dihitung sebagai atribut dengan nilai *gain* tertinggi dibandingkan dengan atribut-atribut lainnya [11]. Persamaan untuk menghitung *gain* adalah sebagai berikut:

$$\text{gain}(S,A) = \text{Entropy} - \sum_{i=1}^n \left(\frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \right) \quad (1)$$

dimana:

- S adalah himpunan kasus.
- A adalah atribut.
- n adalah jumlah partisi atribut A .
- $|S_i|$ adalah jumlah kasus partisi ke- i .
- $|S|$ adalah jumlah kasus dalam S .

Dalam hal memilih atribut untuk setiap *node* pada *tree*, digunakan salah satu metode pengukuran seleksi atribut bernama "gain". Atribut dengan *gain* tertinggi akan dipilih sebagai atribut uji pada setiap *node tree* [12, 13].

Untuk menghitung *gain*, digunakan nilai *entropi* yang dapat dihitung menggunakan persamaan tertentu:

$$\text{Entropy}(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (2)$$

dimana p_i adalah proporsi dari S_i terdapat S .

iii. Naive Bayes

Algoritma Naive Bayes merupakan metode klasifikasi statistik yang digunakan untuk memperkirakan probabilitas keanggotaan suatu kelas [14]. Metode ini berasal dari teorema Bayes dan memiliki kemampuan klasifikasi yang serupa dengan *decision tree* dan *neural network* [15]. Ketika digunakan pada basis data dengan jumlah data yang besar, algoritma Naive Bayes terbukti memiliki tingkat akurasi yang tinggi dan proses yang

cepat [16]. Metode Bayesian merupakan pendekatan statistik yang digunakan untuk inferensi induktif dalam persoalan klasifikasi. Awalnya, penjelasan mengenai konsep dasar dan definisi teorema Bayes dipaparkan, dan setelah itu teorema ini diterapkan untuk melakukan klasifikasi dalam bidang Data Mining [17]. Persamaan dari Naive Bayes adalah

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (3)$$

dimana:

- X adalah data dengan *class* yang belum diketahui.
- H adalah hipotesis data X merupakan suatu *class* spesifik.
- $P(H|X)$ adalah probabilitas hipotesis H berdasarkan kondisi X (*posteriori prob.*).
- $P(H)$ adalah probabilitas hipotesis H (*prior prob.*).
- $P(X|H)$ adalah probabilitas X berdasarkan kondisi tersebut.
- $P(X)$ adalah probabilitas dari X .

iv. Data penelitian

Pengumpulan data pada penelitian ini berupa kuisisioner yang dilakukan kepada konsumen yang pernah membeli tomat dan cabai rawit di pasar Keramat Baamang Sampit. Dari pengumpulan data tersebut diperoleh sebanyak 100 data. Atribut yang digunakan untuk tomat dan cabai rawit adalah karakteristik fisik yang melekat pada kedua jenis produk tersebut. Atribut tomat dan cabai rawit meliputi ukuran, rasa, warna kulit produk, dan tekstur produk dan bentuk produk.

III. HASIL PENELITIAN DAN DISKUSI

i. Preferensi Atribut Tomat dan Cabai Rawit

Dari kelima atribut tomat dan cabai rawit tersebut selanjutnya dikategorikan menjadi lebih spesifik seperti pada Tabel 1 dan 2.

Tabel 1: Atribut Tomat

No.	Atribut	Kategori
1	Rasa	Manis, Asam
2	Warna	Hijau, Kuning, Orange, Merah
3	Tekstur	Padat, Lunak
4	Ukuran	Besar, Sedang, Kecil
5	Bentuk	Bulat pipih, Lonjong

Pada table 1 di atas, atribut tomat dikategorikan menjadi lima kriteria yang lebih spesifik. Kategori atribut tersebut meliputi rasa, warna, tekstur, ukuran, dan

bentuk. Untuk mendapatkan informasi mengenai preferensi konsumen terhadap tomat di pasar keramat baamang sampit, dapat dilihat dari kategori atau kriteria atribut yang mendapatkan pilihan terbanyak dari konsumen. Berdasarkan dari data yang telah dikumpulkan dapat diketahui bahwa tomat yang disukai konsumen dipasar keramat baamang sampit adalah tomat yang mempunyai rasa asam, mempunyai warna merah, ukuran sedang, bentuk nya bulat pipih dan memiliki tekstur yang padat. Selanjutnya ialah mengetahui preferensi konsumen terhadap produk cabai rawit untuk atribut cabai rawit ada di tabel 2 berikut ini.

Tabel 2: Atribut Cabai Rawit

No.	Atribut	Kategori
1	Rasa	Pedas, Kurang pedas
2	Warna	Hijau muda, Kuning, Orange, Merah
3	Tekstur	Cabai yang kulitnya renyah, Cabai yang kulitnya lunak
4	Ukuran	Besar, Sedang, Kecil
5	Bentuk	Bulat panjang ujung meruncing/tumpul, Bulat pendek ujung meruncing/tumpul

Berdasarkan dari data yang telah didapat diketahui bahwa cabai rawit yang disukai konsumen dipasar keramat baamang sampit adalah cabai rawit yang mempunyai rasa pedas, mempunyai warna merah, ukuran sedang, bentuk nya bulat panjang ujung meruncing/tumpul dan memiliki tekstur kulit yang renyah.

ii. Pengolahan Data Tomat: Pengujian Algoritma C45 Menggunakan Rapid Miner

Data yang akan dipakai dalam pengujian ini terdiri dari 100 kuesioner yang telah diberikan kepada para konsumen-konsumen pasar keramat.

Setelah dilakukan penyederhanaan data, langkah selanjutnya adalah memproses data dengan menggunakan aplikasi Rapidminer. Aplikasi ini akan menghasilkan pohon keputusan beserta aturan-aturannya. Berikut adalah hasil pohon keputusan yang diperoleh dari aplikasi tersebut.

Berdasarkan dari pohon keputusan di atas menghasilkan sebuah rule (aturan) melalui aplikasi Rapid Miner. Rule tersebut dapat dilihat pada gambar di bawah ini.

Berdasarkan pada rule yang dihasilkan pada gambar 5, dapat disimpulkan aturan klasifikasi buah berdasarkan atribut seperti ukuran, bentuk, warna, dan tekstur. Buah berukuran besar cenderung memiliki rasa asam (11 data asam dan 3 data manis). Buah berukuran kecil dibagi berdasarkan bentuk, dengan bentuk bulat pipih cenderung memiliki rasa manis (2 data asam dan 3 data manis), dan bentuk lonjong cenderung memiliki

Open in Turbo Prep Auto Model

Row No.	Rasa	Warna	Ukuran	Bentuk	Tekstur
1	Asam	Merah	Besar	Bulat Pipih	Padat
2	Asam	Orange	Sedang	Lonjong	Padat
3	Asam	Orange	Sedang	Bulat Pipih	Lunak
4	Asam	Merah	Besar	Bulat Pipih	Padat
5	Asam	Orange	Sedang	Lonjong	Padat
6	Asam	Orange	Sedang	Bulat Pipih	Padat
7	Asam	Orange	Sedang	Lonjong	Padat
8	Asam	Merah	Sedang	Bulat Pipih	Padat
9	Asam	Merah	Sedang	Lonjong	Padat
10	Asam	Kuning	Kecil	Lonjong	Padat
11	Asam	Kuning	Besar	Lonjong	Padat
12	Asam	Hijau	Sedang	Lonjong	Lunak
13	Manis	Merah	Besar	Bulat Pipih	Padat
14	Asam	Orange	Kecil	Lonjong	Lunak
15	Asam	Merah	Kecil	Bulat Pipih	Lunak

ExampleSet (100 examples, 1 special attribute, 4 regular attributes)

Gambar 2: Penyederhanaan data pada tomat

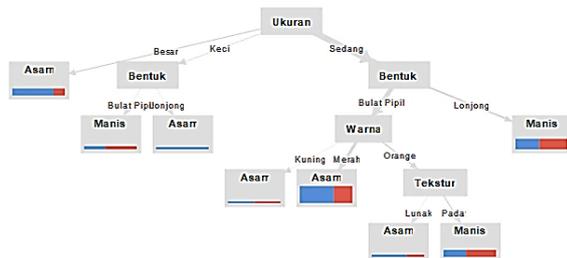
rasa asam (3 data asam dan tidak ada data manis). Buah berukuran sedang memiliki aturan klasifikasi yang lebih kompleks, dengan bentuk bulat pipih dan warna kuning cenderung memiliki rasa asam (1 data asam dan 1 data manis), bentuk bulat pipih dan warna merah cenderung memiliki rasa asam (24 data asam dan 13 data manis), serta bentuk lonjong dan warna orange dengan tekstur lunak cenderung memiliki rasa asam (2 data asam dan 1 data manis), sedangkan dengan tekstur padat cenderung memiliki rasa manis (6 data asam dan 8 data manis). Terakhir, buah berukuran sedang dengan bentuk lonjong cenderung memiliki rasa manis (10 data asam dan 12 data manis). Aturan-aturan ini memberikan panduan yang berguna dalam mengklasifikasikan buah berdasarkan atribut-atribut tertentu untuk menentukan karakteristik rasanya.

Tree

```

Ukuran = Besar: Asam {Asam=11, Manis=3}
Ukuran = Kecil
| Bentuk = Bulat Pipih: Manis {Asam=2, Manis=3}
| Bentuk = Lonjong: Asam {Asam=3, Manis=0}
Ukuran = Sedang
| Bentuk = Bulat Pipih
| | Warna = Kuning: Asam {Asam=1, Manis=1}
| | Warna = Merah: Asam {Asam=24, Manis=13}
| | Warna = Orange
| | | Tekstur = Lunak: Asam {Asam=2, Manis=1}
| | | Tekstur = Padat: Manis {Asam=6, Manis=8}
| Bentuk = Lonjong: Manis {Asam=10, Manis=12}
    
```

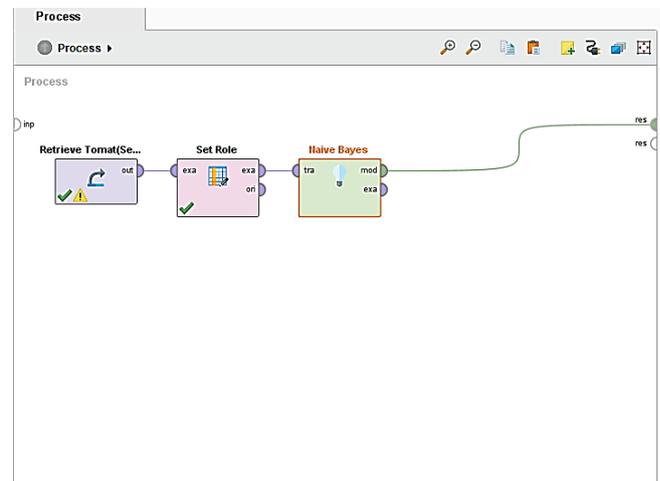
Gambar 4: Rule pada tomat



Gambar 3: Pohon keputusan pada tomat

Pengujian Algoritma Naive Bayes Menggunakan Rapid Miner

Data yang digunakan dalam penelitian ini sama pada gambar nomor 1 yaitu sebanyak 100 data kuesioner.



Gambar 5: Pengujian nilai probabilitas pada tomat

Hasil dari proses pengujian pada Gambar 5 akan menghasilkan probabilitas munculnya setiap nilai dari masing-masing atribut kondisi yang mempengaruhi hasil rasa. Hasil probabilitas yang telah diproses dapat ditemukan dalam Tabel nomor 3.

Tabel 3: Nilai probabilitas pada tomat

Atribut	Parameter	Asam	Manis
Warna	value=Merah	0.609822184589331	0.58489646772229
Warna	value=Orange	0.304995766299746	0.3412911084043849
Warna	value=Kuning	0.06790855207451313	0.0733252131546894
Warna	value=Hijau	0.017104149026248947	2.4360535931790484E-4
Warna	value=unknown	1.6934801016088076E-4	2.4360535931790484E-4
Ukuran	value=Besar	0.18648373983739838	0.0733430799220273
Ukuran	value=Sedang	0.7284891598915989	0.8530701754385964
Ukuran	value=Kecil	0.08485772357723577	0.0733430799220273
Ukuran	value=unknown	1.6937669376693764E-4	2.4366471734892783E-4
Bentuk	value=Bulat Pipih	0.7116720311705912	0.6826712161832805
Bentuk	value=Lonjong	0.2881585634423175	0.3170850597124055
Bentuk	value=unknown	1.694053870913096E-4	2.4372410431391673E-4
Tekstur	value=Padat	0.8980179569710316	0.9020229100658055
Tekstur	value=Lunak	0.10181263764187702	0.09773336582988058
Tekstur	value=unknown	1.694053870913096E-4	2.4372410431391673E-4

Tabel 3 menunjukkan rincian probabilitas untuk atribut dan parameter yang mempengaruhi rasa asam dan manis pada suatu benda. Dalam tabel ini, angka-angka yang diberikan adalah probabilitas yang menunjukkan seberapa besar kemungkinan tomat memiliki rasa asam atau manis berdasarkan atribut dan parameter yang terkait. Apabila diketahui warna tomat merah, ukuran besar, bentuknya bulat pipih, teksturnya padat. Maka perhitungannya adalah:

1. Asam = $0,610 \times 0,186 \times 0,712 \times 0,898 = 0,072$
2. Manis = $0,585 \times 0,073 \times 0,683 \times 0,902 = 0,026$

Dengan melakukan normalisasi terhadap nilai probabilitas di atas, total nilai yang diperoleh dapat dihitung.

$$\text{Probabilitas Asam} = \frac{0.072}{0.072 + 0.026} = 0.7347$$

$$\text{Probabilitas Manis} = \frac{0.026}{0.026 + 0.072} = 0.2653$$

Berdasarkan total jumlah nilai probabilitas yang Asam sebesar 0.7347 dan nilai probabilitas yang manis sebesar 0.2653, dapat disimpulkan bahwa kemungkinan rasa asam pada tomat adalah sekitar 73%, sementara kemungkinan rasa manis adalah sekitar 26%.

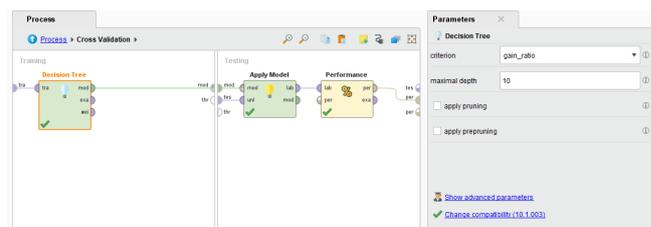
Jadi, rule yang digunakan dalam penentuan klasifikasi adalah sebagai berikut: Jika probabilitas "Asam" lebih besar dari probabilitas "Manis", maka hasilnya adalah "Asam", dan jika probabilitas "Manis" lebih besar dari probabilitas "Asam", maka hasilnya adalah "Manis".

iii. Pengujian Tingkat Akurasi

Dalam pengujian preferensi konsumen menggunakan Rapidminer, operator *Performance* berfungsi untuk

memperoleh nilai akurasi, presisi, dan recall melalui *Confusion Matrix* [18]. *Confusion Matrix* memberikan informasi tentang sejauh mana model klasifikasi dapat memprediksi kelas dengan benar atau salah [19]. Dengan menggunakan metode ini, pengujian data preferensi konsumen memberikan hasil tentang tingkat akurasi model, serta presisi dan recall dari setiap kelas yang dievaluasi.

Algoritma C4.5 dan Naive Bayes diuji untuk mengevaluasi performa klasifikasi pada kelas yang telah ditetapkan. Evaluasi ini dilakukan dengan menggunakan metode K-Fold Validation untuk membagi data menjadi subset yang akan digunakan dalam pengujian. Metode K-Fold Validation membagi data ke dalam K subset atau "lipatan" [8]. Setiap subset akan digunakan sebagai data testing secara bergantian, sementara subset lainnya digunakan sebagai data training [20]. Dalam penelitian ini, akan dilakukan 2 kali pengujian dengan K-Fold Validation yang memiliki nilai yang berbeda. Pertama, dilakukan pengujian dengan K-Fold Validation 10, yang berarti data akan dibagi menjadi 10 subset. Setiap subset akan menjadi data testing satu per satu, sementara 9 subset lainnya digunakan sebagai data training. Kedua, dilakukan pengujian dengan K-Fold Validation 7, yang berarti data akan dibagi menjadi 7 subset. Setiap subset akan menjadi data testing satu per satu, sementara 6 subset lainnya digunakan sebagai data training. Dengan melakukan pengujian sebanyak 2 kali diharapkan dapat mendapatkan informasi yang komprehensif mengenai kinerja klasifikasi algoritma C4.5 (decision tree) dan naive bayes dalam penelitian ini.



Gambar 6: Pengujian menggunakan operator cross validation pada algoritme c45 dengan nilai kfold-10

Pada pengujian pertama, data akan dibagi menjadi 10 bagian, di mana 9 data digunakan untuk training dan 1 data digunakan untuk testing. Algoritma C4.5 digunakan dengan operator Cross Validation untuk melatih data yang telah dibagi. Proses training ini melibatkan pembentukan Decision Tree yang kemudian digunakan untuk menguji data testing menggunakan operator Apply Model dan Performance. Hasil pengujian menunjukkan bahwa nilai akurasi yang diperoleh adalah sebesar 60%.

Pada Tabel 5 menunjukkan hasil evaluasi klasifikasi

Tabel 4: Confusion matrix

	Predicted Class	
	Yes	No
Class = Yes	a (TP/true positive)	b (FN/false negative)
Class = No	c (FP/false positive)	d (TN/true negative)

Tabel 5: Hasil pengujian dengan nilai k-fold 10 pada data preferensi konsumen terhadap tomat menggunakan C4.5 dengan accuracy: 60% pm 20% (micro average: 60%)

Predicted Class	True Class		Class Precision
	Asam	Manis	
Asam	44	25	63.77%
Manis	15	16	51.61%
Class recall: Asam 74.58%, Manis 39.02%			

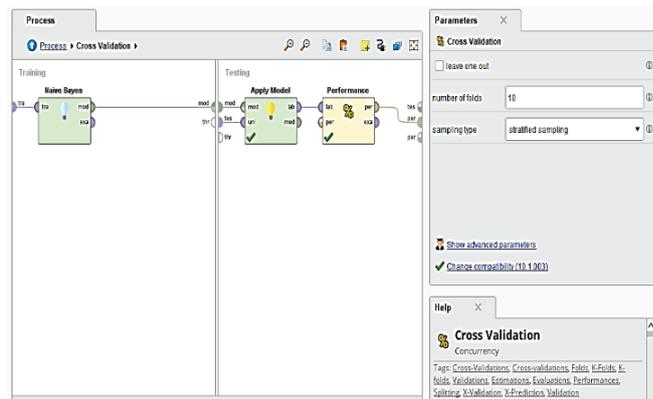
menggunakan aplikasi RapidMiner menyajikan informasi tentang kinerja model dalam memprediksi dua kelas, yaitu "Asam" dan "Manis". Kolom pertama menunjukkan jumlah sampel yang benar-benar termasuk dalam kelas "Asam" (44) dan "Manis" (25). Kolom kedua menunjukkan jumlah sampel yang diprediksi oleh model sebagai "Asam" (15) dan "Manis" (16). Baris pertama memberikan informasi tentang prediksi yang benar untuk kelas "Asam" (44) dan prediksi yang salah untuk kelas "Manis" (16). Baris kedua menampilkan persentase tingkat presisi (precision) untuk kelas "Asam" (63.77%) dan "Manis" (51.61%), menunjukkan akurasi model dalam memprediksi kelas tersebut. Baris ketiga menampilkan persentase tingkat recall untuk kelas "Asam" (74.58%) dan "Manis" (39.02%), menggambarkan kemampuan model dalam mengidentifikasi dengan benar sampel dari setiap kelas. Setelah itu, akan dilakukan pengujian lagi nilai K-fold-7 pada tabel nomor 6 dibawah ini, hasilnya menunjukkan tingkat akurasi dengan nilai 63,06%.

Tabel 6: Hasil pengujian dengan nilai k-fold 7 pada data preferensi konsumen terhadap tomat menggunakan C4.5 dengan accuracy: 63.06% pm 9.46% (micro average: 63%)

Predicted Class	True Class		Class Precision
	Asam	Manis	
Asam	45	23	66.18%
Manis	14	18	56.25%
Class recall: Asam 74.58%, Manis 39.02%			

Selanjutnya pengujian tingkat akurasi pada algor-

titma naive bayes. Pada pengujian 1, dilakukan pengujian dengan k-fold 10 dimana data akan menjadi 10 bagian, 9 data digunakan sebagai data training dan untuk 1 data digunakan sebagai data testing.



Gambar 7: Pengujian menggunakan operator cross validation pada algoritma naive bayes

Data-data yang sudah dibagi akan dilatih menggunakan algoritma naive bayes dengan menggunakan operator Cross Validation. Pada pengujian dengan nilai 10-Fold untuk model Naive Bayes ini hasil akurasinya adalah 45% berikut ini hasil dari pengujiannya.

Tabel 7: Hasil pengujian dengan nilai k-fold 10 pada data preferensi konsumen terhadap tomat menggunakan naive bayes dengan accuracy: 45.00% ± 15.09% (micro average: 45.00%)

Predicted Class	True Class		Class Precision	
	Asam	Manis	Precision	Recall
Pred. Asam	44	40	52.38%	74.58%
Pred. Manis	15	1	6.25%	2.44%

Pada tabel 7 dapat dilihat mengenai hasil evaluasi klasifikasi dengan dua kelas, yaitu "Asam" dan "Manis". Pertama, kolom "True Asam" dan "True Manis" menunjukkan jumlah data yang sebenarnya masuk dalam kelas "Asam" dan "Manis". Selanjutnya, kolom "Pred. Asam" dan "Pred. Manis" menampilkan jumlah data yang diprediksi sebagai "Asam" dan "Manis" oleh model klasifikasi. Kemudian, kolom "Class Precision" menunjukkan presisi klasifikasi untuk setiap kelas, yaitu seberapa akurat model dalam memprediksi data yang

termasuk dalam kelas tersebut. Dalam tabel ini, presisi untuk kelas "Asam" adalah 52.38%, sedangkan untuk kelas "Manis" adalah 6.25%. Terakhir, kolom "Class Recall" menunjukkan recall klasifikasi untuk setiap kelas, yaitu seberapa banyak data aktual dari kelas tersebut yang berhasil diprediksi dengan benar oleh model. Dalam tabel ini, recall untuk kelas "Asam" adalah 74.58%, sementara untuk kelas "Manis" hanya 2.44%. Setelah itu, akan dilakukan pengujian dengan metode K-fold cross-validation dengan nilai 7. Hasil dari pengujian dengan nilai k-fold 7 adalah 49% seperti yang terlihat dalam tabel 8.

Tabel 8: Hasil pengujian K-fold dengan nilai k-fold 7 pada data preferensi konsumen terhadap tomat menggunakan c45 dengan accuracy: 49.05% ± 11.78% (micro average: 49.00%)

Predicted Class	True Class		Class Precision	
	Asam	Manis	Precision	Recall
Pred. Asam	45	38	54.76%	77.97%
Pred. Manis	13	3	18.75%	7.32%

Berdasarkan dari hasil pengujian kedua algoritma masing-masing maka dapat diketahui hasilnya sebagai berikut:

Tabel 9: Hasil perbandingan kedua algoritma

	C45	Naive Bayes
Accuracy 1	60.00%	45.00%
Accuracy 2	63.06%	49.05%

Dari perbandingan kinerja dilakukan, dapat dilihat bahwa Algoritma C4.5 mendapatkan tingkat akurasi sekitar 60,00-63,06%, sementara Algoritma Naive Bayes memiliki tingkat akurasi sekitar 45,00-49,05%. Dari hasil tersebut, dapat dinyatakan bahwa algoritma C4.5 menunjukkan tingkat akurasi yang lebih tinggi dibandingkan dengan Algoritma Naive Bayes. Sehingga, Algoritma C4.5 dapat dianggap memiliki performa yang lebih baik dalam melakukan klasifikasi pada data preferensi konsumen terkait pemilihan tomat.

iv. Pengolahan Data Cabai Rawit: Pengujian Algoritma c45 menggunakan rapid miner

Data yang akan dipakai dalam pengujian ini adalah data kuesioner dimana pengumpulan datanya sama dengan pengumpulan pada data tomat sebelumnya dimana data nya telah disebar kepada para konsumen-konsumen pasar keramat sebanyak 100 kuesioner.

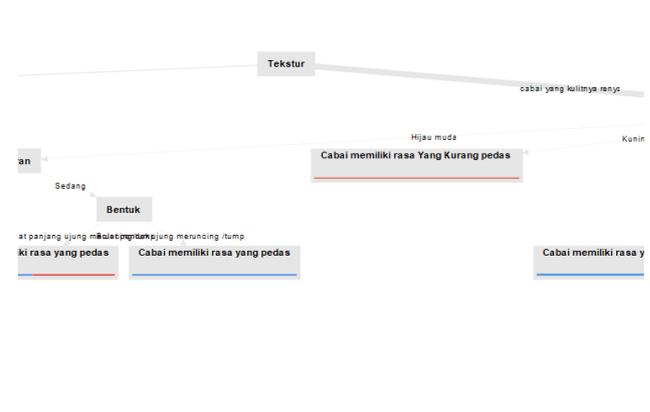
Gambar 9: Data Kuisisioner Prefrensi konsumen terhadap cabai rawit Dari data yang diperoleh menyay-

takan 91 menyukai rasa pedas dan dan yang menyukai rasa kurang pedas 9 orang. Setelah dilakukan penyederhanaan data, langkah selanjutnya adalah memproses data dengan menggunakan aplikasi Rapidminer. Selanjutnya akan menghasilkan pohon keputusan beserta aturan-aturannya. Metode algoritma yang digunakan dalam aplikasi Rapidminer adalah C4.5. Berikut adalah hasil dari pohon keputusan yang diperoleh.

Row No.	Rasa	Warna	Ukuran	Bentuk	Tekstur
1	Cabai memili...	merah	Besar	Bulat panjang...	cabai yang ku...
2	Cabai memili...	merah	Sedang	Bulat pendek...	cabai yang ku...
3	Cabai memili...	Orange	Sedang	Bulat pendek...	cabai yang ku...
4	Cabai memili...	Hijau muda	Sedang	Bulat panjang...	cabai yang ku...
5	Cabai memili...	merah	Sedang	Bulat pendek...	cabai yang ku...
6	Cabai memili...	merah	Sedang	Bulat panjang...	cabai yang ku...
7	Cabai memili...	Orange	Sedang	Bulat panjang...	cabai yang ku...
8	Cabai memili...	merah	Besar	Bulat pendek...	cabai yang ku...
9	Cabai memili...	Orange	Sedang	Bulat pendek...	cabai yang ku...
10	Cabai memili...	Kuning	Kecil	Bulat panjang...	cabai yang ku...
11	Cabai memili...	Hijau muda	Sedang	Bulat panjang...	cabai yang ku...
12	Cabai memili...	Hijau muda	Kecil	Bulat panjang...	cabai yang ku...
13	Cabai memili...	merah	Kecil	Bulat panjang...	cabai yang ku...
14	Cabai memili...	Orange	Besar	Bulat panjang...	cabai yang ku...
15	Cabai memili...	Hijau muda	Sedang	Bulat pendek...	cabai yang ku...

Gambar 8: Data Kuisisioner Prefrensi konsumen terhadap cabai rawit

Dari data yang diperoleh menyatakan 91 menyukai rasa pedas dan dan yang menyukai rasa kurang pedas 9 orang. Setelah dilakukan penyederhanaan data, langkah selanjutnya adalah memproses data dengan menggunakan aplikasi Rapidminer. Selanjutnya akan menghasilkan pohon keputusan beserta aturan-aturannya. Metode algoritma yang digunakan dalam aplikasi Rapidminer adalah C4.5. Berikut adalah hasil dari pohon keputusan yang diperoleh.



Gambar 9: Pohon keputusan pada cabai rawit

Pohon keputusan yang dihasilkan melalui aplikasi Rapidminer akan menghasilkan Rule/aturan yang ada pada gambar 11.

Berdasarkan pohon keputusan dan rule/ aturan yang dihasilkan, dapat ditarik kesimpulan mengenai aturan klasifikasi buah berdasarkan atribut seperti ukuran, bentuk, warna, dan tekstur. Tekstur cabai yang

```

Tree
Tekstur = cabai yang kulitnya lunak
|
|  Ukuran = Besar: Cabai memiliki rasa Yang Kurang pedas (Cabai memiliki rasa yang pedas=0, Cabai memiliki rasa Yang
|  Ukuran = Kecil
|  Warna = Hijau muda: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=1, Cabai memiliki rasa Yang
|  Warna = merah: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=1, Cabai memiliki rasa Yang Kurang
|  Ukuran = Sedang
|  Warna = Orange: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=1, Cabai memiliki rasa Yang Kurang
|  Warna = merah
|  |  Bentuk = Bulat panjang ujung meruncing/tumpul: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang ped
|  |  Bentuk = Bulat pendek ujung meruncing /tumpul: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang ped
Tekstur = cabai yang kulitnya renyah
|  Warna = Hijau muda
|  |  Ukuran = Besar: Cabai memiliki rasa Yang Kurang pedas (Cabai memiliki rasa yang pedas=0, Cabai memiliki rasa Ye
|  |  Ukuran = Sedang
|  |  Bentuk = Bulat panjang ujung meruncing/tumpul: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang ped
|  |  Bentuk = Bulat pendek ujung meruncing /tumpul: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang ped
|  Warna = Kuning: Cabai memiliki rasa Yang Kurang pedas (Cabai memiliki rasa yang pedas=0, Cabai memiliki rasa Yang P
|  Warna = Orange
|  |  Bentuk = Bulat panjang ujung meruncing/tumpul
|  |  Ukuran = Besar: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=3, Cabai memiliki rasa Yang
|  |  Ukuran = Kecil: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=2, Cabai memiliki rasa Yang
|  |  Ukuran = Sedang: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=7, Cabai memiliki rasa Yang
|  |  Bentuk = Bulat pendek ujung meruncing /tumpul: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=
|  Warna = merah
|  |  Bentuk = Bulat panjang ujung meruncing/tumpul
|  |  Ukuran = Besar: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=7, Cabai memiliki rasa Yang
|  |  Ukuran = Kecil: Cabai memiliki rasa yang pedas (Cabai memiliki rasa yang pedas=2, Cabai memiliki rasa Yang
    
```

Gambar 10: Rule pada cabai rawit

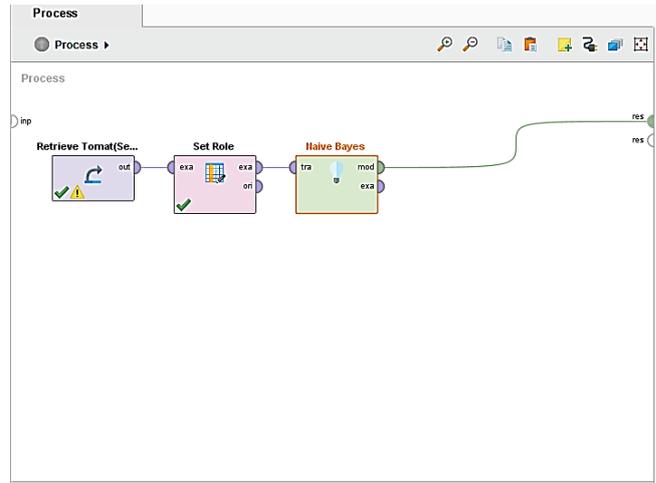
kulitnya lunak memiliki variasi rasa tergantung pada ukuran, warna, dan bentuknya. Jika cabai memiliki ukuran besar, maka cabai tersebut memiliki rasa yang kurang pedas. Namun, jika cabai memiliki ukuran kecil dengan warna hijau muda, cabai tersebut memiliki rasa yang pedas. Selain itu, cabai dengan warna merah juga memiliki rasa yang pedas. Cabai dengan tekstur yang kulitnya renyah juga memiliki variasi rasa tergantung pada warna dan bentuknya. Jika cabai memiliki warna hijau muda dan ukuran kecil, cabai tersebut memiliki rasa yang kurang pedas. Namun, jika cabai memiliki warna kuning, cabai tersebut juga memiliki rasa yang kurang pedas. Selain itu, cabai dengan warna orange memiliki rasa yang pedas, tergantung pada bentuknya. Jika cabai memiliki bentuk bulat panjang dengan ujung meruncing atau tumpul, cabai tersebut memiliki rasa yang pedas. Sedangkan jika cabai memiliki bentuk bulat pendek dengan ujung meruncing atau tumpul, cabai tersebut memiliki rasa yang pedas. Cabai dengan warna merah dan bentuk bulat panjang dengan ujung meruncing atau tumpul juga memiliki variasi rasa tergantung pada ukurannya. Jika cabai memiliki ukuran besar, cabai tersebut memiliki rasa yang pedas. Namun, jika cabai memiliki ukuran kecil atau sedang, cabai tersebut memiliki rasa yang kurang pedas.

v. Pengujian Algoritma naive bayes menggunakan rapid miner

Data yang akan dipakai dalam pengujian ini sama pada gambar nomor 9 yaitu sebanyak 100 kuesioner. Selanjutnya akan diuji menggunakan aplikasi rapid miner. Hasilnya, dapat dilihat dari koneksi antara data dan operator tersebut pada gambar 12.

Dari Proses diatas akan menghasilkan probabilitas kemunculan setiap nilai dari atribut kondisi yang mempengaruhi hasil rasa. Hasil probabilitas yang telah diolah ada pada Tabel 9.

Pada tabel 9 menyajikan rincian probabilitas untuk atribut dan parameter yang mempengaruhi rasa pedas dan kurang pedas pada cabai. Dalam tabel ini, angka-



Gambar 11: Pengujian probilitas pada data preferensi cabai rawit

Tabel 10: Nilai probabilitas cabai rawit

Atribut	Parameter	Cabai Pedas	Cabai Kurang Pedas
Warna	value=merah	0.7689181768259199	0.44309392265193365
Warna	value=Orange	0.1868204283360791	0.2220994475138121
Warna	value=Hijau muda	0.04404173531026909	0.2220994475138121
Warna	value=Kuning	1.0982976386600773E-4	0.11160220994475135
Warna	value=unknown	1.0982976386600773E-4	0.0011049723756906072
Ukuran	value=Besar	0.12093585237258347	0.11172566371681418
Ukuran	value=Sedang	0.8019551845342706	0.5542035398230089
Ukuran	value=Kecil	0.07699912126537783	0.3329646017699115
Ukuran	value=unknown	1.0984182776801412E-4	0.001106194602654871
Bentuk	value=Bulat panjang ujung meruncing/tumpul	0.6702186092496979	0.6655592469545958
Bentuk	value=Bulat pendek ujung meruncing/tumpul	0.3296715368559816	0.3333333333333333
Bentuk	value=unknown	1.0985389432055374E-4	0.0011074197120708748
Tekstur	value=cabai yang kulitnya renyah	0.8899263978908053	0.5548172757475083
Tekstur	value=cabai yang kulitnya lunak	0.10996374821487423	0.4440753045404208
Tekstur	value=unknown	1.0985389432055374E-4	0.0011074197120708748

angka yang diberikan merupakan probabilitas yang menunjukkan seberapa besar kemungkinan cabai memiliki rasa yang pedas atau kurang pedas berdasarkan atribut dan parameter yang terkait. Apabila diketahui warna cabai rawit merah, ukuran besar, bentuknya bulat pipih, teksturnya padat. Maka perhitungannya adalah:

1. Pedas = $0,768 \times 0,120 \times 0,670 \times 0,889 = 0,059$
2. Kurang pedas = $0,443 \times 0,111 \times 0,665 \times 0,554 = 0,017$

Dengan melakukan normalisasi terhadap nilai probabilitas di atas, total nilai yang diperoleh dapat dihitung:

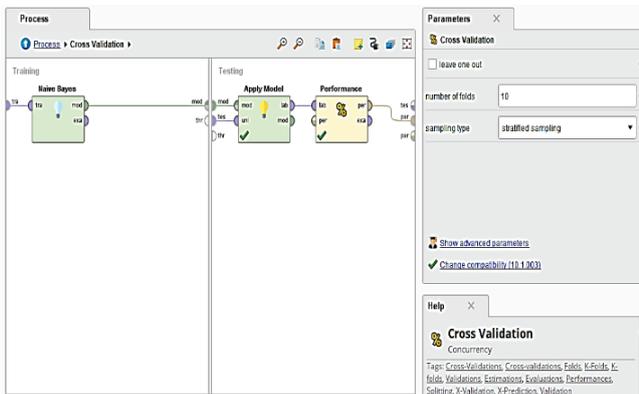
$$\text{Probabilitas Asam} = \frac{0,059}{(0,059 + 0,017)} = 0.775,$$

$$\text{Probabilitas Manis} = \frac{0,017}{(0,017 + 0,059)} = 0.223.$$

Berdasarkan total jumlah nilai probabilitas yang Pedas sebesar 0,775 dan nilai probabilitas yang kurang pedas sebesar 0,223, dapat disimpulkan bahwa kemungkinan rasa pedas pada cabai rawit adalah sekitar 77%, sementara kemungkinan rasa yang kurang pedas adalah sekitar 26%. Jadi perbandingan antara nilai probabilitas "Asam" yang merupakan 0,7347 dan nilai probabilitas "Manis" yang merupakan 0,2653. Maka kesimpulannya

adalah data testing tersebut termasuk dalam klasifikasi "Pedas". Maka rule/aturan yang digunakan dalam penentuan klasifikasi adalah sebagai berikut:

1. Jika probabilitas "Pedas" lebih besar dari probabilitas "Kurang pedas", maka hasilnya adalah "Pedas".
2. Dan jika probabilitas "Kurang pedas" lebih besar dari probabilitas "Pedas", maka hasilnya adalah "Kurang pedas".



Gambar 12: Pengujian tingkat akurasi dengan cross validation menggunakan naive bayes

Tabel 11: Hasil confusion matrix dengan k-fold10 naive bayes pada cabai rawit dengan accuracy: 90.00% ± 6.67% (micro average: 90.00%)

	True Cabai pedas	True Kurang pedas	Class Precision
Pred. Cabai pedas	89	8	91.75%
Pred. Cabai Kurang pedas	2	1	33.33%
Class Recall	97.80%	11.11%	

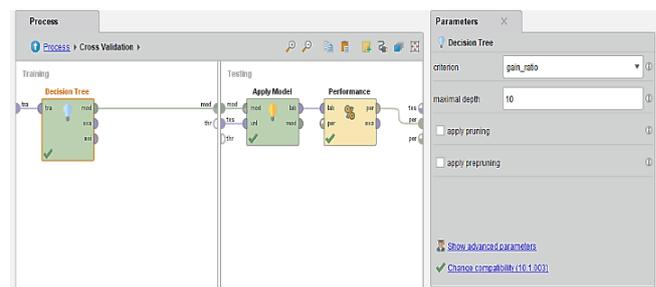
Pada tabel nomor 10 terdiri dari dua kelas, yaitu "Cabai memiliki rasa yang pedas" dan "Cabai memiliki rasa yang kurang pedas", serta ukuran akurasi dan recall untuk setiap kelas. Pada prediksi untuk kelas "Cabai memiliki rasa yang pedas", terdapat 89 prediksi yang benar bahwa cabai tersebut memiliki rasa yang pedas (True Positive), namun juga terdapat 8 prediksi yang salah dimana sebenarnya cabai tersebut memiliki rasa yang pedas namun diprediksi sebagai cabai yang kurang pedas (False Negative). Akurasi prediksi untuk kelas ini adalah 91.75%. Sementara itu, pada prediksi untuk kelas "Cabai memiliki rasa yang kurang pedas", terdapat 2 prediksi yang benar (True Positive) dan 1 prediksi yang salah dimana sebenarnya cabai tersebut memiliki rasa yang kurang pedas namun diprediksi sebagai cabai yang pedas (False Negative). Akurasi prediksi untuk kelas ini adalah 33.33%. Selanjutnya, pada kolom "class recall" atau recall kelas, terlihat bahwa recall untuk kelas "Cabai memiliki rasa yang pedas" mencapai 97.80%, yang menunjukkan bahwa sebagian besar cabai yang memang memiliki rasa yang pedas berhasil

diprediksi dengan benar. Namun, recall untuk kelas "Cabai memiliki rasa yang kurang pedas" hanya sebesar 11.11%, menandakan bahwa sebagian besar cabai yang sebenarnya memiliki rasa yang kurang pedas tidak berhasil diprediksi dengan benar. Kemudian, akan dilakukan pengujian dengan nilai K-7. Hasil validasi dari pengujian ini menunjukkan akurasi sebesar 89

Selanjutnya pengujian tingkat akurasi pada algoritma c45. Pada pengujian 1, akan dilakukan uji dengan dengan k-fold 10. Skenario ini, data dibagi menjadi dua bagian: 9 data digunakan sebagai data training (pelatihan) untuk melatih model, dan 1 data digunakan sebagai data testing (pengujian) untuk menguji kinerja model yang telah dilatih. Setelah melakukan pengujian, diperoleh nilai akurasi sebesar 89%.

Tabel 12: Hasil confusion matrix dengan k-fold10 c45 pada cabai rawit dengan accuracy: 89.00% ± 7.38% (micro average: 89.00%)

	True Cabai pedas	True Cabai Kurang pedas	Class Precision
Pred. Cabai pedas	89	9	90.82%
Pred. Cabai Kurang pedas	2	0	0.00%
Class Recall	97.80%	0.00%	



Gambar 13: Pengujian confusion matrix dengan k-fold 10 pada algoritma c45

Pada tabel 12 menunjukkan hasil evaluasi dari model klasifikasi terhadap prediksi atribut "Cabai" dan rasa yang dimilikinya. Dalam tabel tersebut, terdapat dua kelas yang diamati, yaitu "Cabai memiliki rasa yang pedas" dan "Cabai memiliki rasa yang kurang pedas". Dari hasil evaluasi, dapat dilihat bahwa terdapat 89 prediksi yang benar untuk kelas "Cabai memiliki rasa yang pedas". Artinya, model berhasil mengklasifikasikan sebagian besar data yang benar-benar termasuk ke dalam kelas tersebut. Namun, terdapat pula 9 prediksi yang salah untuk kelas tersebut. Sementara itu, pada kelas "Cabai memiliki rasa yang kurang pedas", terdapat 2 prediksi yang diklasifikasikan sebagai kelas cabai yang pedas tersebut, tetapi tidak ada satupun prediksi yang sesuai dengan kelas yang sebenarnya. Oleh karena itu, nilai precision untuk kelas ini adalah 0.00%, menunjukkan bahwa model tidak efektif dalam mengklasifikasikan data ke dalam kelas tersebut. Selain itu dapat dilihat nilai recall untuk kedua kelas. Recall menggambarkan

seberapa baik model dalam mendeteksi data yang benar-benar termasuk dalam suatu kelas. Pada kelas "Cabai memiliki rasa yang pedas", model berhasil mendeteksi sekitar 97.80% dari keseluruhan data yang sebenarnya termasuk dalam kelas tersebut. Namun, untuk kelas "Cabai memiliki rasa yang kurang pedas", model tidak berhasil mendeteksi satupun data yang termasuk ke dalam kelas tersebut, sehingga recall-nya adalah 0.00%. Selanjutnya adalah hasil dari pengujian dengan k-fold 7 adalah sebesar 89% berikut hasil dari pengujiannya.

Tabel 13: Hasil confusion matrix dengan k-fold7 c45 pada cabai rawit dengan nilai accuracy: 89.05% \pm 3.58% (micro average: 89.00%)

	True Cabai pedas	True Cabai Kurang pedas	Class Precision
Pred. Cabai pedas	89	9	90.82%
Pred. Cabai Kurang pedas	2	0	0.00%
Class Recall	97.80%	0.00%	

Setelah melakukan analisis dari kedua pengujian algoritma yang telah dilakukan, berikut adalah temuan dan hasilnya dapat dilihat pada tabel 14:

Tabel 14: Hasil perbandingan akurasi kedua algoritma

	C45	Naive Bayes
Accuracy 1	89.00%	90.00%
Accuracy 2	89.05%	89.05%

Berdasarkan perbandingan kinerja yang telah diuraikan, dapat dilihat bahwa Algoritma C45 dan Algoritma Naive Bayes menghasilkan akurasi yang berbeda. Algoritma C45 mencapai akurasi sekitar 89,00-89,05%, sementara Algoritma Naive Bayes mencapai akurasi sekitar 89,05-90,00%. Dari hasil tersebut, terlihat bahwa Algoritma Naive Bayes memiliki tingkat akurasi yang lebih tinggi daripada Algoritma C45 dalam memprediksi preferensi konsumen terkait rasa cabai rawit. Hasil dari perbandingan ini menunjukkan bahwa Algoritma Naive Bayes lebih baik dalam melakukan klasifikasi pada data preferensi konsumen dalam memilih cabai rawit.

IV. KESIMPULAN

Berdasarkan pembahasan sebelumnya, telah diuraikan tentang dua metode yang digunakan dalam analisis preferensi konsumen dalam memilih tomat dan cabai rawit di pasar keramat baamang sampit, yaitu menggunakan algoritma Naive Bayes dan C4.5. Dari hasil analisis dan pengujian yang telah dilakukan dapat disimpulkan konsumen yang menyukai tomat dan cabai rawit dapat di klasifikasikan untuk tomat konsumen lebih suka warna merah, ukurannya sedang, bentuknya

bulat pipih, memiliki tekstur padat, dan memiliki cita rasa asam. Sedangkan untuk cabai rawit konsumen menyukai warna yang merah, ukuran sedang, bentuknya bulat panjang ujung meruncing/tumpul dan memiliki tekstur kulit yang renyah. Selanjutnya dari hasil pengujian menggunakan kedua metode algoritma c45 dan naive bayes maka dapat disimpulkan: Dalam penelitian ini dari hasil terhadap analisis preferensi konsumen terhadap tomat, kedua algoritma, yaitu Naive Bayes dan C45, diimplementasikan menggunakan software RapidMiner. Berdasarkan hasil pengujian, algoritma C4.5 menunjukkan tingkat akurasi antara 60,00% hingga 63,06%, sementara algoritma Naive Bayes memiliki tingkat akurasi sekitar 45,00% hingga 49,05%. Dari perbandingan ini, dapat di tarik kesimpulan bahwa algoritma C45 mempunyai performa yang lebih baik dalam melakukan klasifikasi pada data preferensi konsumen terhadap tomat dalam penelitian ini. Dalam penelitian ini, analisis preferensi konsumen terhadap cabai rawit, kedua algoritma yaitu Naive Bayes dan C45, melalui software RapidMiner. Berdasarkan perbandingan kinerja, algoritma C45 menunjukkan tingkat akurasi antara 89,00% hingga 89,05%, sedangkan algoritma Naive Bayes memiliki tingkat akurasi sekitar 89,05% hingga 90,00%. Dari perbandingan tersebut, dapat di tarik kesimpulan bahwa algoritma Naive Bayes mempunyai performa yang lebih baik dalam melakukan klasifikasi pada data preferensi konsumen terhadap rasa cabai rawit dalam penelitian ini.

DAFTAR PUSTAKA

- [1] S. Wahyuni, M. Aqiel, N. Ismi, S. Mahdianti, U. Usman, and F. Fachrian, "Ameliorasi lahan berpasir sebagai cara pemulihan lahan berpasir menjadi sawah produktif di kebun percobaan pt pim." *J. Serambi Eng.*, vol. 8, no. 1, pp. 4689-4696, 2023.
- [2] S. Wigati, "Perilaku konsumen dalam prespektif ekonomi islam," *Maliyah J. Huk. Bisnis Islam*, vol. 01, no. 01, p. 18, 2011.
- [3] M. W. Herdian, R. Winanjaya, and Susiani, "Penerapan algoritma c4.5 dalam klasifikasi kualitas sayur kol di kabupaten simalungun," *Resolusi Rekayasa Tek. Inform. dan Inf.*, vol. 2, no. 3, pp. 132-137, 2022. [Online]. Available: <https://doi.org/10.30865/resolusi.v2i3.312>
- [4] M. S. Firdaus, A. Primajaya, and A. Jamaludin, "Implementasi algoritme c4.5 untuk prediksi penanaman cabai merah," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 6, no. 2, p. 158, 2021. [Online]. Available: <https://doi.org/10.30998/string.v6i2.10378>
- [5] A. Nainggolan, H. Rumapea, A. P. Silalahi, L. Sidauruk, and M. Sinambela, "Identifikasi penyakit tanaman tomat berdasarkan citra penyakit menggunakan metode glem dan naive bayes classifier." *J. Ilm. Tek. Inform.*, vol. 2, no. 1, pp. 22-28, 2022. [Online]. Available: <http://ojs.fikom-methodist.net/index.php/METHOTIKA>
- [6] Styawati, A. Nurkholis, Z. Abidin, and H. Sulistiani, "Optimasi parameter support vector machine berbasis

- algoritma firefly pada data opini film,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 5, pp. 904–910, 2021. [Online]. Available: <https://doi.org/10.29207/resti.v5i5.3380>
- [7] M. I. Putri and I. Kharisudin, “Penerapan synthetic minority oversampling technique (smote) terhadap analisis sentimen data review pengguna aplikasi marketplace tokopedia,” *Prism. Pros. Semin. Nas. Mat.*, vol. 5, pp. 759–766, 2022. [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [8] U. Riyanto, “Analisis perbandingan algoritma naive bayes dan support vector machine dalam mengklasifikasikan jumlah pembaca artikel online,” *JIKA (Jurnal Inform.)*, vol. 2, no. 2, pp. 62–72, 2019. [Online]. Available: <https://doi.org/10.31000/v2i2.1521>
- [9] B. Novianti, T. Rismawan, and S. Bahri, “Implementasi data mining dengan algoritma c4.5 untuk penjurusan siswa (studi kasus: Sma negeri 1 pontianak),” *J. Coding, Sist. Komput. Untan*, vol. 04, no. 3, pp. 75–84, 2016.
- [10] F. D. M. Achmad, Budanis, and Slamant, “Klasifikasi data karyawan untuk menentukan jadwal kerja menggunakan metode decision tree,” *J. IPTEK*, vol. 16, no. 1, pp. 18–23, 2012. [Online]. Available: <http://jurnal.itats.ac.id/wp-content/uploads/2013/06/3.-BUDANIS-FINAL-hal-17-23.pdf>
- [11] H. P. Tambunan and S. Zetli, “Jurnal comasie,” *Comasie*, vol. 3, no. 3, pp. 21–30, 2020.
- [12] A. Essra, Rahmadani, and Safriadi, “Analisis information gain attribute evaluation untuk klasifikasi serangan intrusi,” *J. Inf. Syst. Dev.*, vol. 2, no. 2, pp. 9–14, 2016.
- [13] A. E. Wijaya, R. Bani, S. Sukarni, and S. A. Weighting, “Jurnal teknologi informasi dan komunikasi stmik subang,” *no. April*, pp. 100–110, 2019.
- [14] O. Arifin and T. B. Sasongko, “Analisa perbandingan tingkat performansi metode support vector machine dan naive bayes classifier,” *Semin. Nas. Teknol. Inf. dan Multimed.* 2018, vol. 6, no. 1, pp. 67–72, 2018.
- [15] S. Dewi, “Komparasi 5 metode algoritma klasifikasi data mining pada prediksi keberhasilan pemasaran produk layanan perbankan,” *Techno Nusa Mandiri*, vol. 13, no. 1, pp. 60–66, 2016.
- [16] H. Pramudia and A. Nugroho, “Sistem informasi kerusakan laptop menggunakan metode naive bayes,” *Teknol. Elektro, Univ. Mercu Buana*, vol. 8, no. 3, pp. 206–214, 2017. [Online]. Available: <https://publikasi.mercubuana.ac.id/files/journals/4/articles/2186/submission/original/2186-4652-1-SM.pdf>
- [17] H. Annur, “Klasifikasi masyarakat miskin menggunakan metode naive bayes,” *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018. [Online]. Available: <https://doi.org/10.33096/ilkom.v10i2.303.160-165>
- [18] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, “Analisis sentimen ulasan villa di ubud menggunakan metode naive bayes, decision tree, dan k-nn,” *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 3, pp. 205–215, 2022. [Online]. Available: <https://doi.org/10.23887/janapati.v11i3.49450>
- [19] E. Prasetyaningrum and P. Susanti, “Analisa tingkat kepuasan pelanggan pada percetakan cv. mega media menggunakan algoritma c4.5,” *Sisfotenika*, vol. 13, no. 1, pp. 65–75, 2023.
- [20] R. Tuntun, K. Kusriani, and K. Kusnawi, “Analisis perbandingan kinerja algoritma klasifikasi dengan menggunakan metode k-fold cross validation,” *J. Media Inform. Budi-darma*, vol. 6, no. 4, p. 2111, 2022. [Online]. Available: <https://doi.org/10.30865/mib.v6i4.4681>