

ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ID3 UNTUK FAKTOR KEPUASAN KONSUMEN WARUNG ICHA STEAK & SEAFOOD

Muhammad Farhan Arief¹, Ir Herry Wahyono², Nuke L Chusna³

^{1,2,3}Prodi Teknik Informatika Universitas Krisnadwipayana, Jakarta
farhanarief94@gmail.com, wahyonos2000@unkris.ac.id, nukelchusna@unkris.ac.id

ABSTRACT- Warung Icha Steak & Seafood is an *MSME* type food stall business that focuses on the culinary field that sells food in the form of steak and processed food from the sea. Efforts to retain consumers are certainly not an easy thing, especially the thing that is now being faced by Icha Steak & Seafood stalls is declining the number of consumers and customers. This is a threat that is being faced by the *msme* because it will have an impact on revenue turnover. The purpose of this study is to find out what factors need to be improved to achieve consumer satisfaction of Warung Icha Steak & Seafood. Apply the Classification method with C4.5 and ID3 algorithms to group the most influential factors for consumer satisfaction. Perform a Comparison of both C4.5 and ID3 algorithms with the Cross Validation method. It was concluded that the problem of determining the consumer satisfaction factor of Warung Icha Steak & Seafood can be solved using data mining techniques, namely with the C4.5 and ID3 algorithms, resulting in the same decision tree and 6 *rules* and the most dominant factor is Service Friendliness (C8) with a gain value of 0.41370089. Testing the ID3 algorithm has a high level of accuracy of 98.50%, a Precision value of 98.77% and a *Recall* value of 99.38%.

Keywords : Customer Satisfaction, Comparison Analysis, C4.5 Algorithm, ID3 Algorithm

1. PENDAHULUAN

Latar Belakang

Kepuasan konsumen menjadi hal yang paling penting bagi para pelaku (UMKM) Usaha Mikro Kecil dan Menengah dibidang kuliner. Didalam dunia kuliner. Kepuasan konsumen adalah suatu keadaan yang dihadirkan konsumen pada saat ia menyadari bahwa kebutuhan dan keinginannya sesuai dengan harapannya atau tidak. Ketika tingkat kepuasan tinggi, ada kemungkinan untuk membeli kembali produk tersebut sehingga pendapatan pelaku di sektor tersebut tetap stabil atau bahkan meningkat secara proporsional (Novita Indriyani et al., 2022).

Icha Steak & Seafood merupakan usaha warung makanan berjenis UMKM yang berfokus didalam bidang kuliner yang menjual makanan berupa steak dan makanan olahan dari laut. Warung Icha Steak & Seafood 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 warung Icha Steak & Seafood yaitu sedang menurunnya jumlah konsumen dan pelanggan warung Icha Steak & Seafood. Hal tersebut menjadi sebuah ancaman yang sedang dihadapi oleh toko tersebut karena akan berdampak pada omset pendapatan.

Namun untuk meningkatkan kepuasan konsumen warung Icha Steak & Seafood belum dapat mengetahui aspek-aspek apa saja yang harus di tingkatkan dikarenakan dengan sistem keluhan konsumen yang hanya disediakan contact person oleh Icha Steak & Seafood. Dengan demikian, kepuasan konsumen sangat sulit diketahui, karena tidak semua konsumen menghubungi contact person yang disediakan.

Beberapa faktor yang mempengaruhi kepuasan konsumen terhadap produk yaitu : Harga, Rasa, Kebersihan tempat, Variasi Produk, Kecepatan Penyajian, Lamanya Antrian Orderan, Ketersediaan Lahan Parkir

dan Keramahan Pelayanan. Dengan demikian, pihak UMKM harus membuat kebijakan untuk meningkatkan jumlah konsumen kembali dan kemudian mencari faktor yang paling dominan dalam mempengaruhi kepuasan konsumen.

Dalam mengetahui faktor yang paling berpengaruh 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 (Novita Indriyani et al., 2022). Salah satu Algoritma klasifikasi yang bisa digunakan adalah C4.5 dan ID3.

Cross Validasi digunakan untuk menguji keefektifan model pembelajaran mesin dengan teknik validasi dengan membagi seluruh data menjadi dua set pelatihan dan set tes, model k-Fold Cross Validation digunakan untuk mengetahui nilai k Optimal yang melibatkan pemisahan secara acak kumpulan sampel menjadi serangkaian fold (grup) berukuran sama, di mana k menunjukkan jumlah partisi atau fold dari dataset yang dibagi menjadi data latih dan data uji, Misalnya, jika nilai-k sepuluh digunakan, maka dataset dibagi menjadi sepuluh partisi. Dalam hal ini, sembilan dari partisi digunakan untuk data latih, sedangkan satu partisi sisanya digunakan untuk data uji (Tanjung, 2021).

Dengan menggunakan metode Algoritma C4.5 dan ID3, penelitian ini diharapkan mampu mengetahui apa yang perlu ditingkatkan untuk mencapai kepuasan dari pelanggan, serta untuk membantu pemilik Warung Icah Steak & Seafood dalam menganalisa dan mengevaluasi kinerja karyawan sehingga bisa dijadikan model bahan pengambilan keputusan di masa yang akan datang.

Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mencari faktor-faktor apa saja yang perlu ditingkatkan untuk mencapai kepuasan

konsumen Warung Icah Steak & Seafood. Menerapkan metode Klasifikasi dengan Algoritma C4.5 dan ID3 untuk mengelompokkan faktor-faktor yang paling berpengaruh untuk kepuasan konsumen. Melakukan perbandingan kedua Algoritma C4.5 dan ID3 dengan metode Cross Validation.

2. LANDASAN TEORI

Data Mining

Data mining ialah tahap penganalisan yang digunakan untuk menganalisis kondisi pasar dan memprediksi jumlah produksi yang tepat bagi sebuah produk dari database (Lesmana & Silalahi, 2020). Data mining disebut juga dengan Knowledge Discovery in Database (KDD) ataupun pattern recognition digunakan untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Sedangkan istilah pattern recognition atau disebut pengenalan pola mempunyai tujuan pengetahuan yang akan digali dari dalam bongkahan data yang sedang dihadapi (Batubara & Windarto, 2019). 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 (Novita Indriyani et al., 2022).

Cross-Industry Standard Process for Data Mining (CRISP-DM)

Data mining telah diterapkan di hamper seluruh bidang industri dan pengetahuan. Dengan semakin luasnya penerapan data mining tersebut, terdapat keinginan dari sekelompok analis data mining yang mewakili Daimler Chrysler, SPSS, dan NCR untuk membuat sebuah model proses data mining yang netral terhadap jenis industri, tools, dan aplikasi. Para analis tersebut bergabung dan membangun Cross-Industry Standard Process for Data mining (CRISP-

DM) pada tahun 1996. CRISP-DM menyediakan proses standar pelaksanaan data mining untuk menyelesaikan masalah dalam sebuah bisnis atau penelitian yang dapat digunakan semua orang (Rahmayuni, 2014).

Algoritma C4.5

Algoritma ini merupakan penerus dari Algoritma sebelumnya yaitu Algoritma ID3 digunakan untuk menghasilkan aturan dalam bentuk pohon keputusan. Algoritma ini dapat digunakan untuk klasifikasi data dan disebut classifier 30eknik30ic. Algoritma C4.5 merupakan Algoritma pengurutan atau pengelompokan yang bersifat prediktif. Pernyataan klasifikasi dalam pohon keputusan berada di cabang dan kelas atau segmen berada di daun. Algoritma C4.5 adalah keluarga pohon keputusan yang dapat menghasilkan pohon keputusan dan aturan untuk meningkatkan akurasi prediksi. Selain itu, model C4.5 mudah digunakan dan dipahami sebagai aturan turunan dan memiliki 30eknik interpretasi yang sangat sederhana (Novita Indriyani et al., 2022). Secara umum Algoritma C4.5 langkah untuk membangun pohon keputusan adalah sebagai berikut (Shiddiq et al., 2018) :

- a) Pilih atribut sebagai akar.
- b) Buat cabang untuk masing-masing nilai.
- c) Bagi kasus dalam cabang.
- d) Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain yang digunakan rumus seperti Persamaan 2.1 yang tertera berikut :

$$Gain(S, A) = Entropy(S) - \sum_i^n = 1 \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.1)$$

Keterangan :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus pada S

Sebelum mendapatkan nilai Gain adalah dengan mencari nilai Entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar Entropi menggunakan sebagai berikut :

$$Entropy(S, A) = - \sum_i^n = 1 P_i * \log_2 P_i$$

Keterangan :

S : Himpunan Kasus

A : Fitur

n : Jumlah partisi S

P_i : Proporsi dari S_i terhadap S

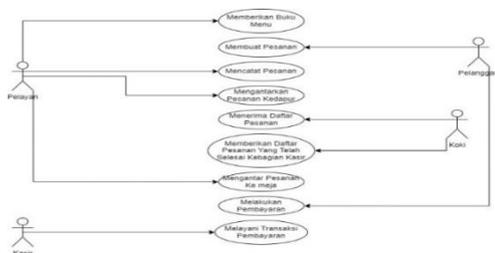
Algoritma ID3

ID3 adalah Algoritma pembelajaran pohon keputusan sederhana yang dikembangkan oleh Ross Quinlan (1983). Pada tahun 1975, ID3 pertama kali dipresentasikan dalam sebuah buku Quinlan. Algoritma Concept Learning System (CLS) adalah dasar dari Algoritma ID3. Dengan menambahkan pemilihan fitur heuristik ID3 untuk meningkatkan CLS. Atribut dari instance pelatihan adalah dicari melalui ID3 dan atribut yang paling baik memisahkan contoh yang diberikan diekstraksi olehnya. ID3 berhenti jika atribut dengan sempurna mengklasifikasikan set pelatihan; jika tidak, secara rekursif beroperasi pada n, di mana n adalah jumlah kemungkinan nilai atribut dari himpunan bagian yang dipartisi untuk mendapatkan atribut "terbaik" mereka. Selain itu, Algoritma ID3 memilih atribut sebagai informasi entropy yang akan dihitung berdasarkan probabilitas, dan perolehan informasi untuk mengembangkan pohon keputusan (Tanjung, 2021) Untuk membangun pohon keputusan, hitung information gain untuk setiap atribut dan pilih atribut dengan information gain tertinggi untuk dijadikan sebagai akar. Beri label pada atribut sebagai akar. Information Gain merupakan suatu ukuran korelasi pada

model parametrik yang menggambarkan ketergantungan antara dua peubah acak X dan Y. Terdapat beberapa tahapan dalam membentuk pohon keputusan dengan Algoritma ID3 (Melina Agustina, 2016).

Cross Validation

Cross Validation adalah teknik untuk memvalidasi akurasi model yang dibangun pada kumpulan data tertentu, yang membagi kumpulan data menjadi dua bagian, yaitu data pelatihan dan data pengujian. Untuk masalah prediksi, model biasanya diberi dataset dari data yang diketahui untuk dilatih (dataset pelatihan) dan data yang tidak diketahui (atau data yang pertama kali muncul) untuk menguji model disebut validasi atau data uji (Destiningrum & Adrian, 2017). Cross Validation merupakan suatu metode tambahan dari teknik data mining pengembangan model split validation yang bertujuan untuk memperoleh hasil akurasi yang maksimal dimana validasinya mengukur training error dengan menguji dengan test data atau data uji (Tanjung, 2021).



Confusion Matrix

Confusion Matrix sering digunakan sebagai metode untuk mengukur kinerja pada metode pengelompokkan. Kerja dari metode ini dilakukan dengan membandingkan data yang sudah diuji menjadi ada empat kategori istilah yang disematkan untuk menggambarkan hasil proses yakni Positif Salah (TN), Negatif Salah, Positif Benar (TP), dan Negatif Benar (TN) (Hijrah, 2022). Evaluasi dengan confusion matrix menghasilkan nilai accuracy, precision, dan recall. Accuracy dalam klasifikasi adalah persentase ketepatan record data yang diklasifikasikan secara benar setelah dilakukan pengujian

pada hasil klasifikasi, (Han & Kamber, 2006). Sedangkan precision atau confidence adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. Recall atau sensitivity adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar,(Gali et al., 2019).

3. ANALISA

Identifikasi Masalah

Berdasarkan latar belakang yang ada dapat di indentifikasi masalah sebagai berikut:

1. Faktor apa saja yang harus ditingkatkan warung icha & steak untuk mencapai kepuasan konsumen ?
2. Bagaimana cara Algoritma C4.5 dan ID3 menentukan faktor yang paling mempengaruhi untuk kepuasan kepuasan warung icha steak & seafood?
3. Bagaimana hasil perbandingan akurasi, presisi dan *recall* Algoritma C4.5 dan ID3 dengan metode Cross Validation untuk data kepuasan konsumen?

Analisis Sistem Berjalan

Analisis sistem berjalan merupakan gambaran mengenai sistem yang ada saat ini ada di Warung Icha Steak & Seafood.

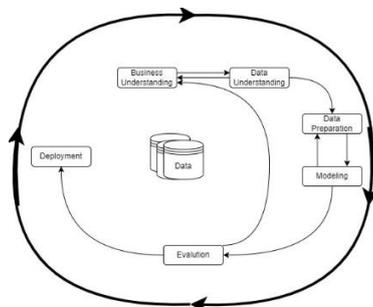
Gambar 1. Usecase Pemesanan Makanan

1. Pelayan memberikan buku menu kepada konsumen untuk konsumen memilih makanan atau minuman yang ingin dipesan.
2. Konsumen memberi tahu makanan atau minuman yang ingin dimakan kepada pelayan.
3. Pelayan mencatat pesanan yang diberikan oleh konsumen.
4. Pelayan mengantarkan daftar pesanan konsumen kebagian dapur.
5. Koki menerima daftar pesanan & membuat makanan yang di pesan konsumen.
6. Koki memberikan daftar pesanan yang telah selesai kebagian kasir.

7. Pelayan mengantarkan makanan yang selesai dibuat ke meja konsumen.
8. Konsumen melakukan pembayaran dikasir.
9. Kasir melayani transaksi pembayaran konsumen.

Metode Crips DM

Penelitian yang digunakan adalah penelitian Experiment. Penelitian eksperimen melibatkan penyelidikan perlakuan parameter atau variabel tergantung pada penelitiannya dan menggunakan tes yang dikendalikan oleh si peneliti itu sendiri.



Gambar 2. Proses Metode Crips DM

A. Bussines Understanding

Berdasarkan data kepuasan konsumen warung icha steak & seafood yang telah dikumpulkan dari bulan November 2022, terlihat masih banyaknya konsumen yang tidak puas berjumlah 39 orang dari data yang dikumpulkan sebanyak 200 data, hal ini menjadi sebuah masalah dan kendala bagi warung icha steak & seafood karena akan berdampak pada omset pemasukan yang akan berkurang. Dalam beberapa penelitian telah dilakukan proses analisa kepuasan konsumen dengan Algoritma C4.5 dan ID3, dengan tujuan analisa yang dilakukan lebih akurat.

B. Data Understanding

Data yang didapatkan dari warung icha steak & seafood adalah data yang diperoleh dari hasil kuesioner yang telah diberikan untuk diisi oleh konsumen Warung Icha Steak & Seafood. Data responden dikumpulkan melalui pembagian kuesioner melalui berkas

survei yang telah dicetak untuk Warung Icha Steak & Seafood. Jumlah responden yang digunakan adalah 200 responden. Atribut atau variable yang ada sebanyak 11 atribut (data lengkapnya bisa dilihat dilampiran). Setelah dilakukan proses data *preparation*, atribut atau variable yang digunakan terdiri dari 8 atribut yang terdapat dalam data kepuasan konsumen. Variable predictor yaitu Harga, Rasa, Kebersihan Tempat, Variasi Produk, Kecepatan Penyajian, Lamanya Antrian Orderan, Ketersediaan Lahan Parkir, Keramahan Pelayanan. Variabel responnya adalah kepuasan yaitu SP (Sangat Puas), P (Puas), CP (Cukup puas), TP (Tidak puas). Sedangkan variable tujuan adalah kepuasan konsumen (yang menunjukkan puas atau tidak puas).

C. Data Preparation

Pada tahapan ini data sebanyak 200 dan atribut yang terdiri dari 11 atribut, akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan, tahapannya yaitu :

1. Data *Cleaning* untuk membersihkan nilai yang kosong atau tuple yang kosong.
2. Data *Intergration* yang berfungsi menyatukan tempat penyimpanan yang berbeda kedalam satu data.
3. Data *Reduction*, Jumlah atribut yang digunakan terlalu banyak dan tidak semua atribut menjadi syarat atas atribut penentu sehingga yang digunakan hanya 8 atribut.

D. Modeling

Pada tahapan ini, data *preparation* yang telah didapatkan akan dilakukan permodelan Algoritma C4.5, dimana dalam probabilitas itu dilakukan perhitungan manual, kemudian dimasukan ke dalam tools *RapidMiner* dengan menggunakan Algoritma C4.5.

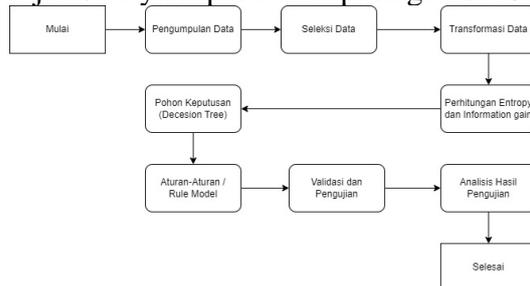
E. Evaluasi

Pada tahapan ini dilakukan untuk mengetahui apakah pola atau informasi yang diketahui bertentangan dengan fakta yang ada sebelumnya. Pada tahap evaluasi akan dilakukan pengujian terhadap data testing dengan metode Cross Validation.

Evaluasi dilakukan dengan melihat informasi dari Confusion Matrix untuk mengetahui tingkat keakuratan prediksi yang diperoleh dari Algoritma yang digunakan.

Model Analisis

Pada bagian model analisis berisi diagram alir metode yang terdiri dari preprocessing data Algoritma C4.5. berikut penjelasannya dapat dilihat pada gambar 3.3.



Gambar 3. Diagram Alir Metode

Pada gambar 3.3 proses klasifikasi dimulai dari pengumpulan data mentah yang didapatkan dari pengisian kuesioner yang telah diisi oleh konsumen Warung Icha Steak & Seafood data mentah tersebut kemudian dimasukkan ke dalam aplikasi excel.

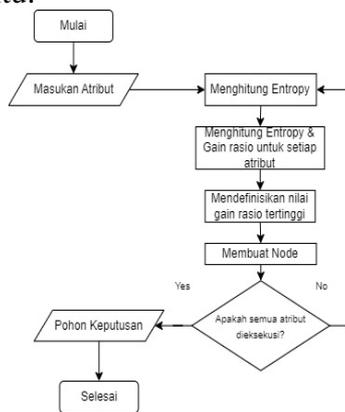
Tahap kedua yaitu Seleksi data adalah memilih data yang akan digunakan dalam proses Algoritma klasifikasi C4.5. Tujuan dari seleksi data adalah menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (discovery) akan dilakukan. Tahap ketiga yaitu transformasi data yang di dapat dari hasil pengisian kuesioner yang telah diisi oleh konsumen diubah menjadi angka/numerik sesuai dengan nilai kepuasan yang terdiri dari SP (Sangat Puas) bernilai 4, P (Puas) bernilai 3, CP (Cukup puas) bernilai 2, TP (Tidak puas) bernilai 1. Tahap keempat perhitungan semua atribut/variabel, entropy menggunakan rumus dan information gain menggunakan rumus untuk mengetahui information gain tertinggi yang akan dijadikan simpul akar pada pembuatan pohon keputusan. Tahap kelima adalah

membuat pohon keputusan adalah hasil dari proses perhitungan entropy dan information gain, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan. Tahap keenam adalah aturan-aturan / rule model adalah uraian penjelasan yang merepresentasikan sebuah pohon keputusan. Tahap ketujuh adalah validasi dan pengujian adalah Pengujian yang dilakukan untuk mengetahui semua fungsi bekerja dengan baik atau tidak. Validasi dilakukan dengan Ten-fold Cross Validation. Ten-fold Cross Validation adalah validasi yang dilakukan dengan cara membagi suatu set data menjadi sepuluh segmen yang berukuran sama besar dengan cara melakukan pengacakan data. Validasi dan pengujian dilakukan untuk mengetahui tingkat akurasi, presisi, dan recall dari hasil prediksi klasifikasi. Akurasi adalah persentase dari catatan yang diklasifikasikan dengan benar dalam pengujian dataset. Presisi adalah persentase data yang diklasifikasikan sebagai model baik yang sebenarnya juga baik. Recall adalah pengukuran tingkat pengenalan positif sebenarnya dan Tahap terakhir adalah Analisis hasil pengujian Analisa yang dilakukan untuk memastikan bahwa hasil pengujian benar-benar sesuai dengan pembahasan. Analisa dilakukan dengan melakukan perhitungan kembali hasil validasi dan pengujian (akurasi, presisi, dan recall) secara manual. Apakah perhitungan yang dilakukan akan menghasilkan nilai yang sama atau tidak, dibantu dengan Confusion matrix. Confusion Matrix adalah model yang akan membentuk matrix yang terdiri dari true positif atau tupel positif dan true negatif atau tupel negatif. Confusion matrix berisi informasi aktual (actual) dan prediksi (predicted) pada sistem klasifikasi. Confusion matrix dapat memvisualisasi kinerja Algoritma klasifikasi.

Proses Perhitungan Manual Algoritma C4.5

Setelah semua data kepuasan konsumen Warung Icha Steak & Seafood

ditransformasi ke dalam bentuk angka/numerik maka data tersebut dapat dikelompokkan dengan menggunakan Algoritma C4.5. Untuk dapat melakukan pengelompokan data-data tersebut untuk membuat pohon keputusan maka perlu dilakukan langkah-langkah seperti gambar 3.4, yaitu:



Gambar 4. Flowchart 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 2.1 . Selanjutnya dilakukan perhitungan mencari nilai Gain untuk setiap atribut dengan Persamaan 2.2 .

Perhitungan Manual Node 1

a) Menghitung entropy total :

$$\text{Entropy [Total]} = \left(-\frac{161}{200} \times \log_2 \frac{161}{200}\right) + \left(-\frac{39}{200} \times \log_2 \frac{39}{200}\right) = 0,71181467$$

b) Menghitung entropy dan gain Harga :

$$\text{Entropy [Harga - Sangat Puas]} = \left(-\frac{96}{102} \times \log_2 \frac{96}{102}\right) + \left(-\frac{6}{102} \times \log_2 \frac{6}{102}\right) = 0,322756959$$

$$\text{Entropy [Harga - Puas]} =$$

$$\left(-\frac{65}{88} \times \log_2 \frac{65}{88}\right) + \left(-\frac{23}{88} \times \log_2 \frac{23}{88}\right) = 0,828797155$$

$$\text{Entropy [Harga - Cukup Puas]} = \left(-\frac{0}{10} \times \log_2 \frac{0}{10}\right) + \left(-\frac{10}{10} \times \log_2 \frac{10}{10}\right) = 0$$

$$\text{Entropy [Harga - Tidak Puas]} = \left(-\frac{0}{0} \times \log_2 \frac{0}{0}\right) + \left(-\frac{0}{0} \times \log_2 \frac{0}{0}\right) = 0$$

$$\begin{aligned} \text{Gain [Harga]} &= 0,71181467 - \left(\frac{102}{200} \times 0,322756959\right) \\ &\quad + \left(\frac{88}{200} \times 0,828797155\right) \\ &\quad + \left(\frac{10}{200} \times 0\right) + \left(\frac{0}{200} \times 0\right) \\ &= 0,18253787 \end{aligned}$$

c) Menghitung entropy dan gain Rasa :

$$\text{Entropy [Rasa - Sangat Puas]} = \left(-\frac{89}{95} \times \log_2 \frac{89}{95}\right) + \left(\frac{6}{95} \times \log_2 \frac{6}{95}\right) = 0,339855078$$

$$\text{Entropy [Rasa - Puas]} = \left(-\frac{72}{89} \times \log_2 \frac{72}{89}\right) + \left(-\frac{17}{89} \times \log_2 \frac{17}{89}\right) = 0,703582101$$

$$\text{Entropy [Rasa - Cukup Puas]} = \left(-\frac{0}{16} \times \log_2 \frac{0}{16}\right) + \left(-\frac{16}{16} \times \log_2 \frac{16}{16}\right) = 0$$

$$\text{Entropy [Rasa - Tidak Puas]} = \left(-\frac{0}{0} \times \log_2 \frac{0}{0}\right) + \left(-\frac{0}{0} \times \log_2 \frac{0}{0}\right) = 0$$

$$\begin{aligned} \text{Gain [Rasa]} &= 0,71181467 - \left(\frac{95}{200} \times 0,339855078\right) \\ &\quad + \left(\frac{89}{200} \times 0,703582101\right) + \left(\frac{16}{200} \times 0\right) \\ &\quad + \left(\frac{0}{200} \times 0\right) = 0,237289473 \end{aligned}$$

d) Menghitung entropy dan gain Kebersihan :

$$\text{Entropy [Kebersihan - Sangat Puas]} = \left(-\frac{55}{55} \times \log_2 \frac{55}{55}\right) + \left(-\frac{0}{55} \times \log_2 \frac{0}{55}\right) = 0$$

$$\text{Entropy [Kebersihan - Puas]} = \left(-\frac{85}{87} \times \log_2 \frac{85}{87}\right) + \left(-\frac{2}{87} \times \log_2 \frac{2}{87}\right) = 0,157906374$$

$$\text{Entropy [Kebersihan - Cukup Puas]} =$$

$$\left(-\frac{13}{36} \times \log_2 \left(\frac{13}{36}\right)\right) + \left(-\frac{23}{36} \times \log_2 \left(\frac{23}{36}\right)\right) = 0,943601631$$

$$\text{Entropy [Kebersihan – Tidak Puas]} = \left(-\frac{8}{22} \times \log_2 \left(\frac{8}{22}\right)\right) + \left(-\frac{14}{22} \times \log_2 \left(\frac{14}{22}\right)\right) = 0,945660305$$

$$\begin{aligned} \text{Gain [Kebersihan]} &= 0,71181467 - \left(\left(\frac{55}{200} \times 0\right) + \left(\frac{87}{200} \times 0,157906374\right) + \left(\frac{36}{200} \times 0,943601631\right) + \left(\frac{22}{200} \times 0,945660305\right)\right) \\ &= 0,36925447 \end{aligned}$$

e) Menghitung entropy dan gain Variasi Produk

$$\text{Entropy [Variasi Produk – Sangat Puas]} = \left(-\frac{89}{89} \times \log_2 \left(\frac{89}{89}\right)\right) + \left(-\frac{0}{89} \times \log_2 \left(\frac{0}{89}\right)\right) = 0$$

$$\text{Entropy [Variasi Produk – Puas]} = \left(-\frac{62}{74} \times \log_2 \left(\frac{64}{72}\right)\right) + \left(-\frac{12}{74} \times \log_2 \left(\frac{12}{74}\right)\right) = 0,639457132$$

$$\text{Entropy [Variasi Produk – Cukup Puas]} = \left(-\frac{10}{28} \times \log_2 \left(\frac{10}{28}\right)\right) + \left(-\frac{18}{28} \times \log_2 \left(\frac{18}{28}\right)\right) = 0,940285959$$

$$\text{Entropy [Variasi Produk – Tidak Puas]} = \left(-\frac{0}{9} \times \log_2 \left(\frac{0}{9}\right)\right) + \left(-\frac{9}{9} \times \log_2 \left(\frac{9}{9}\right)\right) = 0$$

$$\begin{aligned} \text{Gain [Variasi Produk]} &= 0,71181467 - \left(\left(\frac{89}{200} \times 0\right) + \left(\frac{74}{200} \times 0,639457132\right) + \left(\frac{28}{200} \times 0,940285959\right) + \left(\frac{9}{200} \times 0\right)\right) \\ &= 0,369306513 \end{aligned}$$

a) Menghitung entropy dan gain Lamanya Antrian

$$\text{Entropy [Lamanya Antrian Orderan – Sangat Puas]} = \left(-\frac{61}{61} \times \log_2 \left(\frac{61}{61}\right)\right) + \left(-\frac{0}{61} \times \log_2 \left(\frac{0}{61}\right)\right) = 0$$

$$\text{Entropy [Lamanya Antrian Orderan – Puas]} = \left(-\frac{69}{69} \times \log_2 \left(\frac{69}{69}\right)\right) + \left(-\frac{0}{69} \times \log_2 \left(\frac{0}{69}\right)\right) = 0$$

$$\begin{aligned} \text{Entropy [Lamanya Antrian Orderan – Cukup Puas]} &= \left(-\frac{25}{57} \times \log_2 \left(\frac{25}{57}\right)\right) + \left(-\frac{32}{57} \times \log_2 \left(\frac{32}{57}\right)\right) = 0,989093440 \end{aligned}$$

$$\text{Entropy [Lamanya Antrian Orderan – Tidak Puas]} =$$

$$\left(-\frac{6}{13} \times \log_2 \left(\frac{6}{13}\right)\right) + \left(-\frac{7}{13} \times \log_2 \left(\frac{7}{13}\right)\right) = 0,995727452$$

$$\begin{aligned} \text{Gain [Lamanya Antrian Orderan]} &= 0,71181467 - \left(\left(\frac{61}{200} \times 0\right) + \left(\frac{69}{200} \times 0\right) + \left(\frac{57}{200} \times 0,989093440\right) + \left(\frac{13}{200} \times 0,995727452\right)\right) \\ &= 0,365200756 \end{aligned}$$

b) Menghitung entropy dan gain Kecepatan Penyajian

$$\text{Entropy [Kecepatan Penyajian – Sangat Puas]} = \left(-\frac{35}{35} \times \log_2 \left(\frac{35}{35}\right)\right) + \left(-\frac{0}{35} \times \log_2 \left(\frac{0}{35}\right)\right) = 0$$

$$\text{Entropy [Kecepatan Penyajian – Puas]} = \left(-\frac{88}{88} \times \log_2 \left(\frac{88}{88}\right)\right) + \left(-\frac{0}{88} \times \log_2 \left(\frac{0}{88}\right)\right) = 0$$

$$\text{Entropy [Kecepatan Penyajian – Cukup Puas]} = \left(-\frac{32}{63} \times \log_2 \left(\frac{32}{63}\right)\right) + \left(-\frac{31}{63} \times \log_2 \left(\frac{31}{63}\right)\right) = 0,999818247$$

$$\text{Entropy [Kecepatan Penyajian – Tidak Puas]} = \left(-\frac{6}{14} \times \log_2 \left(\frac{6}{14}\right)\right) + \left(-\frac{8}{14} \times \log_2 \left(\frac{8}{14}\right)\right) = 0,985228136$$

$$\begin{aligned} \text{Gain [Kecepatan Penyajian]} &= 0,71181467 - \left(\left(\frac{35}{200} \times 0\right) + \left(\frac{88}{200} \times 0\right) + \left(\frac{63}{200} \times 0,999818247\right) + \left(\frac{14}{200} \times 0,985228136\right)\right) \\ &= 0,327905953 \end{aligned}$$

c) Menghitung entropy dan gain Ketersediaan Lahan Parkir

$$\begin{aligned} \text{Entropy [Ketersediaan Lahan Parkir – Sangat Puas]} &= \left(-\frac{29}{29} \times \log_2 \left(\frac{29}{29}\right)\right) + \left(-\frac{0}{29} \times \log_2 \left(\frac{0}{29}\right)\right) = 0 \end{aligned}$$

$$\text{Entropy [Ketersediaan Lahan Parkir – Puas]} = \left(-\frac{71}{71} \times \log_2 \left(\frac{71}{71}\right)\right) + \left(-\frac{0}{71} \times \log_2 \left(\frac{0}{71}\right)\right) = 0$$

$$\begin{aligned} \text{Entropy [Ketersediaan Lahan Parkir – Cukup Puas]} &= \left(-\frac{57}{81} \times \log_2 \left(\frac{57}{81}\right)\right) + \left(-\frac{24}{81} \times \log_2 \left(\frac{24}{81}\right)\right) = 0,876716289 \end{aligned}$$

$$\begin{aligned} \text{Entropy [Ketersediaan Lahan Parkir – Tidak Puas]} &= \left(-\frac{4}{19} \times \log_2 \left(\frac{4}{19}\right)\right) + \left(-\frac{15}{19} \times \log_2 \left(\frac{15}{19}\right)\right) = 0,74248757 \end{aligned}$$

$$\text{Gain [Ketersediaan Lahan Parkir]} =$$

$$0,71181467 - \left(\left(\frac{29}{200} \times 0 \right) + \left(\frac{71}{200} \times 0 \right) + \left(\frac{81}{200} \times 0,876716289 \right) + \left(\frac{19}{200} \times 0,74248757 \right) \right) = 0,286208254$$

d) Menghitung entropy dan gain Keramahan Pelayanan :

$$\text{Entropy [Keramahan Pelayanan – Sangat Puas]} = \left(-\frac{117}{117} \times \log_2 \left(\frac{117}{117} \right) \right) + \left(-\frac{0}{117} \times \log_2 \left(\frac{0}{117} \right) \right) = 0$$

$$\text{Entropy [Keramahan Pelayanan – Puas]} = \left(-\frac{34}{45} \times \log_2 \left(\frac{34}{45} \right) \right) + \left(-\frac{11}{45} \times \log_2 \left(\frac{11}{45} \right) \right) = 0,802353443$$

$$\text{Entropy [Keramahan Pelayanan – Cukup Puas]} = \left(-\frac{10}{24} \times \log_2 \left(\frac{10}{24} \right) \right) + \left(-\frac{14}{24} \times \log_2 \left(\frac{14}{24} \right) \right) = 0,979868757$$

$$\text{Entropy [Keramahan Pelayanan – Tidak Puas]} = \left(-\frac{14}{14} \times \log_2 \left(\frac{14}{14} \right) \right) + \left(-\frac{0}{14} \times \log_2 \left(\frac{0}{14} \right) \right) = 0$$

$$\text{Gain [Keramahan Pelayanan]} = 0,71181467 - \left(\left(\frac{117}{200} \times 0 \right) + \left(\frac{45}{200} \times 0,802353443 \right) + \left(\frac{24}{200} \times 0,979868757 \right) + \left(\frac{14}{200} \times 0 \right) \right) = 0,413700895$$

4. HASIL DAN PEMBAHASAN

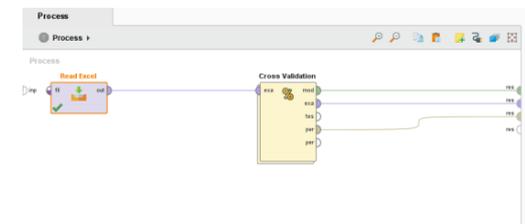
Pengujian Cross Validation

Data diuji dan dianalisis dengan dua Algoritma klasifikasi yaitu Algoritma ID3 dan Algoritma C4.5 (menggunakan cross validasi). Data yang digunakan adalah data kepuasan konsumen yang diambil untuk analisis penelitian ini.

Tabel 1. Dataset

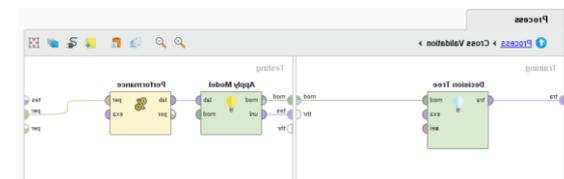
NO	UMUR	SEX	STATUS	STATUS							
1	21	Laki-Laki	Puas	Sangat Puas	Tidak Puas	Sangat Puas	Sangat Puas	Sangat Puas	Puas	Sangat Puas	Puas
2	21	Laki-Laki	Sangat Puas	Sangat Puas	Cukup Puas	Puas	Sangat Puas	Sangat Puas	Cukup Puas	Sangat Puas	Puas
3	18	Perempuan	Sangat Puas	Sangat Puas	Tidak Puas	Sangat Puas	Sangat Puas	Puas	Puas	Puas	Puas
4	20	Perempuan	Puas	Sangat Puas	Sangat Puas	Sangat Puas	Puas	Cukup Puas	Cukup Puas	Sangat Puas	Puas
5	18	Perempuan	Sangat Puas	Puas							
6	22	Laki-Laki	Sangat Puas	Sangat Puas	Sangat Puas	Sangat Puas	Puas	Puas	Puas	Sangat Puas	Puas
7	20	Perempuan	Puas	Sangat Puas	Puas	Sangat Puas	Sangat Puas	Puas	Puas	Sangat Puas	Puas
8	18	Perempuan	Puas	Puas	Puas	Sangat Puas	Cukup Puas	Cukup Puas	Puas	Puas	Puas
9	22	Laki-Laki	Sangat Puas	Puas	Puas	Sangat Puas	Puas	Puas	Cukup Puas	Sangat Puas	Puas
10	20	Perempuan	Sangat Puas	Puas	Tidak Puas	Tidak Puas	Cukup Puas	Tidak Puas	Tidak Puas	Puas	Tidak Puas
11	18	Perempuan	Puas	Puas							
12	20	Laki-Laki	Cukup Puas	Tidak Puas	Tidak Puas						
13	18	Perempuan	Sangat Puas	Puas	Sangat Puas	Sangat Puas	Puas				
14	19	Laki-Laki	Puas	Puas							
15	21	Laki-Laki	Puas	Sangat Puas	Tidak Puas	Cukup Puas	Cukup Puas	Cukup Puas	Tidak Puas	Tidak Puas	Tidak Puas

Kemudian dilakukan transformasi data untuk tahap selanjutnya, hasil dari transformasi data tersebut akan diuji menggunakan desain model Algoritma ID3 dan C4.5 dengan RapidMiner. Pada gambar 4.1 merupakan model yang dirancang di RapidMiner.



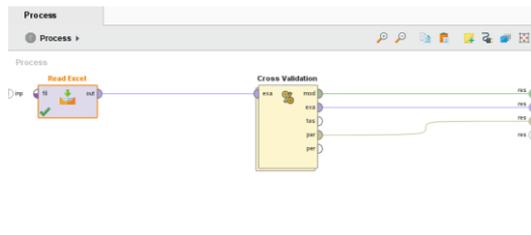
Gambar 5. Model Cross Validation Algoritma C4.5

Data input pada model diatas disimpan dalam format xls dan ada juga set data pelatihan yang diberikan agar sistem dapat belajar dari data yang diberikan.



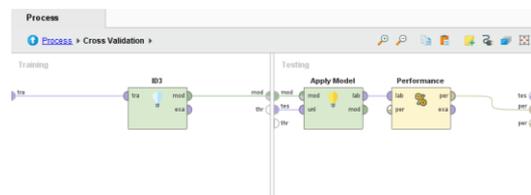
Gambar 6. Pengujian Cross Validation C4.5

Pada gambar 4.2 merupakan model dari Algoritma ID3 yang diterapkan dalam RapidMiner, dimana operator dan parameter yang digunakan akan berpengaruh terhadap akurasi dan model yang terbentuk.



Gambar 7. Model Cross Validation Algoritma ID3

Pada gambar 4.3 diatas sama halnya pada Algoritma sebelumnya yaitu Algoritma C4.5, dimana dataset dihubungkan dengan operator cross validation yang didalamnya terdapat proses seperti gambar berikut.



Gambar 8. Pengujian Cross Validation Algoritma ID3

Pada gambar 4.4 merupakan model pengujian dengan cross validasi, dimana proses dalam membangun model C4.5 dan ID3 dengan menggunakan operator dan parameter. Model yang digunakan juga akan berpengaruh terhadap akurasi yang dihasilkan.

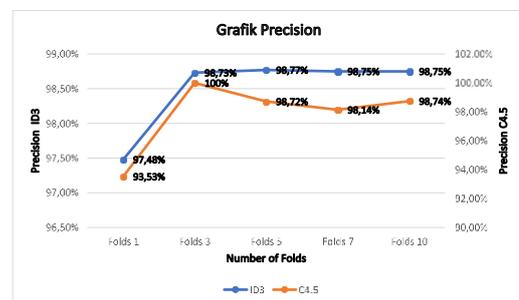
Selanjutnya tahap pengujian yang dilakukan menggunakan model K-Fold Cross Validation untuk menguji tingkat akurasi dari model Algoritma klasifikasi yang digunakan. Cross-validasi merupakan sebuah teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independent (Tanjung, 2021). K-fold cross validation merupakan salah satu teknik dari cross validasi yang digunakan untuk menghilangkan bias pada data. Pelatihan dan pengujian dilakukan sebanyak k kali (Tanjung, 2021). Nilai k-fold cross validation yang digunakan adalah 1, 3, 5, 7 dan 10-fold.

Berikut merupakan peroses perhitungan tabel confusion matrix dari setiap folds untuk Algoritma ID3 dan C4.5 :

Tabel 2. Nilai Presisi perbandingan Algoritma dengan Cross Validation

Algorithm Classification	Number of Fold	Precision
ID3	1	97.48%
	3	98.73%
	5	98.77%
	7	98.75%
	10	98.75%
C4.5	1	93.53%
	3	100%
	5	98.72%
	7	98.14%
	10	98.74%

Perbandingan dari kedua Algoritma klasifikasi dengan dengan nilai presisi ditunjukkan dalam grafik yang terdapat pada gambar 4.5.



Gambar 9. Grafik Precision

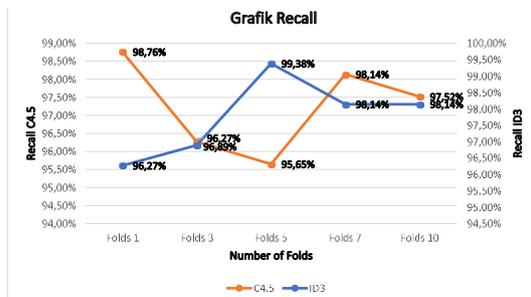
Hasil pengujian dengan model cross validation menunjukkan bahwa Algoritma C4.5 memiliki nilai precision 100%, sedangkan pada Algoritma ID3 nilai precision tertinggi sebesar 98,77% dengan nilai fold = 5. Sehingga Algoritma C4.5 menunjukkan nilai precision lebih tinggi dari pada Algoritma ID3.

Sedangkan hasil *Recall* Algoritma ID3 dan C4.5 dapat dilihat pada tabel 3 berikut :

Tabel 3. Nilai *Recall* perbandingan Algoritma dengan Cross Validation

Algorithm Classification	Number of Fold	Recall
ID3	1	96.27%
	3	96.89%
	5	99.38%
	7	98.14%
	10	98.14%
C4.5	1	98.76%
	3	96.27%
	5	95.65%
	7	98.14%
	10	97.52%

Perbandingan dari kedua Algoritma klasifikasi dengan dengan nilai Recall ditunjukkan dalam grafik yang terdapat pada gambar 4.6.



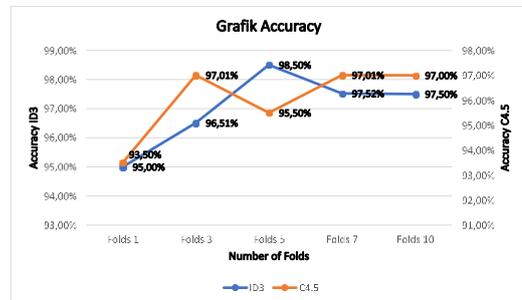
Gambar 10. Grafik Recall

Hasil pengujian dengan model cross validation menunjukkan bahwa Algoritma ID3 memiliki nilai recall 99,38%, sedangkan pada Algoritma C4.5 nilai recall tertinggi sebesar 98,76% dengan nilai fold = 1. Sehingga Algoritma ID3 menunjukkan nilai recall lebih tinggi dari pada Algoritma C4.5. Selanjutnya merupakan hasil akurasi dari perbandingan Algoritma ID3 dan C4.5 yang dapat dilihat pada tabel 4. berikut :

Tabel 4. Nilai Accuracy perbandingan Algoritma dengan Cross Validation

Algorithm Classification	Number of Fold	Accuracy
ID3	1	95.00%
	3	96.51%
	5	98.50%
	7	97.52%
	10	97.50%
C4.5	1	93.50%
	3	97.01%
	5	95.50%
	7	97.01%
	10	97.00%

Perbandingan dari kedua Algoritma klasifikasi dengan dengan nilai Accuracy ditunjukkan dalam grafik yang terdapat pada gambar 11.



Gambar 11 Grafik Accuracy

Dari hasil pengujian validasi dengan beberapa nilai fold dapat diketahui bahwa Algoritma ID3 memiliki nilai accuracy lebih tinggi dibandingkan dengan Algoritma C4.5 dengan nilai akurasi 98,50% sedangkan Algoritma C4.5 dengan nilai 97,01% pada nilai fold = 3 dan nilai fold = 7. Dilihat dari hasil k-fold yang lain juga menunjukkan hasil accuracy dari Algoritma ID3 lebih tinggi daripada Algoritma C4.5.

Oleh sebab itu, dapat disimpulkan bahwa Algoritma ID3 memiliki tingkat akurasi yang lebih tinggi berdasarkan pengujian jumlah data. Sehingga untuk membentuk pohon keputusan berdasarkan hasil akurasi tertinggi yaitu Algoritma ID3. Dimana hasil pohon keputusan dapat dijadikan acuan untuk prediksi kepuasan konsumen kedepannya yang dapat dilihat pada gambar 12.



Gambar 12. Pohon Keputusan ID3

Rule yang dihasilkan dari pohon keputusan pada gambar 12 menggunakan tools *RapidMiner* adalah sebagai berikut :

C8(Keramahan Pelayanan) = Cukup Puas

| C6 (Kecepatan Penyajian) = Cukup Puas: Tidak Puas {Puas=0, Tidak Puas=12}
 | C6 (Kecepatan Penyajian) = Puas: Puas {Puas=7, Tidak Puas=0}
 | C6 (Kecepatan Penyajian) = Sangat Puas: Puas {Puas=3, Tidak Puas=0}
 | C6 (Kecepatan Penyajian) = Tidak Puas: Tidak Puas {Puas=0, Tidak Puas=2}
 C8(Keramahan Pelayanan) = Puas
 | C5(Lamanya Antrian Orderan) = Cukup Puas
 | | C4(Variasi Produk) = Sangat Puas: Puas {Puas=3, Tidak Puas=0}
 | | C4(Variasi Produk) = Tidak Puas: Tidak Puas {Puas=0, Tidak Puas=6}
 | C5(Lamanya Antrian Orderan) = Puas: Puas {Puas=13, Tidak Puas=0}
 | C5(Lamanya Antrian Orderan) = Sangat Puas: Puas {Puas=18, Tidak Puas=0}
 | C5(Lamanya Antrian Orderan) = Tidak Puas: Tidak Puas {Puas=0, Tidak Puas=5}
 C8(Keramahan Pelayanan) = Sangat Puas: Puas {Puas=117, Tidak Puas=0}
 C8(Keramahan Pelayanan) = Tidak Puas: Tidak Puas {Puas=0, Tidak Puas=14}

5. KESIMPULAN

Berdasarkan seluruh hasil tahapan penelitian yang telah dilakukan pada Penerapan Algoritma C4.5 pada Analisis Perbandingan Algoritma C4.5 dan ID3 Untuk Faktor Kepuasan Konsumen Warung Icha Steak & Seafood dapat disimpulkan :

1. Faktor-Faktor yang harus di tingkatkan untuk mencapai kepuasan konsumen antara lain adalah : Kecepatan Penyajian, Lamanya Antrian Orderan , Variasi Produk dan didapatkan faktor yang paling dominan adalah Keramahan Pelayanan (C8) dengan nilai gain 0,41370089 sebagai akar dari pohon keputusan.
2. Algoritma C4.5 dan ID3 dapat menentukan faktor kepuasan konsumen Warung Icha Steak & Seafood dengan mencari gain dan entropy dari setiap atribut dan didapatkan faktor yang paling dominan adalah Keramahan Pelayanan (C8) dengan nilai gain 0,413700895 sebagai akar dari pohon keputusan, menghasilkan pohon keputusan yang sama dengan 6 rules.
3. Berdasarkan pengujian dengan metode Cross Validation terhadap Algoritma ID3 dan C45 didapatkan hasil Algoritma ID3 memiliki tingkat akurasi yang lebih tinggi yaitu sebesar 98,50%,

nilai Precision sebesar 98,77% dan nilai Recall yaitu sebesar 99,38%.

DAFTAR PUSTAKA

- Bangsa, S. T., & Utara, S. (2020). *Analisa Kepuasan Pelanggan Menggunakan Klasifikasi Data Mining*. 2(1), 41–48.
- Batubara, D. N., & Windarto, A. P. (2019). Analisa Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung Taman Hewan Pematang Siantar Dengan Algoritma. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 588–592. <https://doi.org/10.30865/komik.v3i1.1664>
- Destiningrum, M., & Adrian, Q. J. (2017). Jurnal Teknoinfo. Universitas Teknokrat Indonesia, 11(2), 30-37. <http://ejournal.pelitanusantara.ac.id/index.php/man tik/article/view/25/20>
- Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi. *Computer Engineering, Science and System Journal*, 3(2), 173. <https://doi.org/10.24114/cess.v3i2.9661>
- Gali, N., Mariescu-Istodor, R., Hostettler, D., & Fränti, P. (2019). Framework for syntactic string similarity measures. *Expert Systems with Applications*, 129, 169–185. <https://doi.org/10.1016/j.eswa.2019.03.048>
- Hijrah, H. (2022). Analisis Perbandingan Aplikasi Data Mining Dalam Memprediksi Kualitas Kinerja Karyawan Menggunakan Metode Algoritma C4.5. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 1655–1665. <https://doi.org/10.35957/jatisi.v9i2.1992>
- Kamil, M., & Cholil, W. (2020). Analisis Perbandingan Algoritma C4.5 dan Naive Bayes pada Lulusan Tepat Waktu Mahasiswa di Universitas Islam Negeri Raden Fatah Palembang. *Jurnal Informatika*, 7(2), 97–106. <https://doi.org/10.31294/ji.v7i2.7723>
- Melina Agustina, D. (2016). Analisis Perbandingan Algoritma ID3 Dan C4.5 Untuk Klasifikasi Penerima Hibah Pemasangan Air Minum Pada PDAM Kabupaten Kendal Comparative Analysis Of ID3 And C4.5 Algorithm For Classification Of Grant Recipients Of Drinking Water Installation At PDAM Kend. *Journal of Applied Intelligent System*, 1(3), 234–244.
- Novita Indriyani, Heru Satria Tambunan, & Zulia Almaida Siregar. (2022). Analisis Faktor Kepuasan Konsumen Terhadap Produk Roti Pinkan Bakery & Cake dengan Algoritma C4.5. *Jural Riset Rumpun*

Ilmu Teknik, 1(2), 76–90.
<https://doi.org/10.55606/jurritek.v1i2.413>

Putri, S. M., & Arnomo, S. A. (2020). Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus: Hinet Batam). *Journal of Information System Research (JOSH)*, 1(2), 70–76. <http://ejournal.seminar-id.com/index.php/josh/article/view/69>

Rahmayuni, I. (2014). Perbandingan Performansi Algoritma C4.5 dan Cart Dalam Klasifikasi Data Nilai Mahasiswa Prodi Teknik Komputer Politeknik Negeri Padang. *Teknoif*, 2(1), 40–46.

Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 192. <https://doi.org/10.24114/cess.v5i2.18519>

Shiddiq, A., Niswatin, R. K., & Farida, I. N. (2018). Ahmad Shiddiq Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri). *Generation Journal*, 2(1), 9. <https://doi.org/10.29407/gj.v2i1.12051>

Tanjung, D. Y. H. (2021). Analisis perbandingan algoritma id3 dan c4.5 terhadap data pengisian uang atm. *CSRID Journal*, 13(3A), 231–242.

Tree, M., & Algoritma, C. (2018). Decision Tree Klasifikasi-Decision Tree Data Mining Pemilihan Atribut Entropi adalah nilai informasi yang menyatakan Decision Tree • Sebuah Decision Tree adalah struktur yang dapat digunakan untuk membagi data yang besar menjadi himpunan-himpunan record y. 1–5.

Widianto, K., Rifai, A., & Pratiwi, M. N. (2019). Analisa Kepuasan Pengunjung Menggunakan Metoddecision Treedi Perkampungan Budaya Betawi Setu Babakan Jakarta. *Jurnal Mantik Penusa*, 3(23), 301–316.