

Penerapan Data Mining Klasifikasi Kepuasan Pelanggan Transportasi *Online* Menggunakan Algoritma C4.5

Nenitrolina Ndruru¹, Anita Sindar²

^{1,2}Teknik Informatika, STMIK Pelita Nusantara
Jl. Iskandar Muda No. 1 Medan

¹⁾sayaneni1998ndr@gmail.com, ²⁾haito_ita@yahoo.com

Abstract-The classification of customer satisfaction of online transportation users is a positive attitude or customer response to the use of online transportation services. This classification of online transportation user satisfaction is carried out with the aim of improving the quality of the company, especially in the field of service for two-wheeled transportation users. In this study, the level of customer satisfaction is classified into four attributes, namely price, facilities, service, and loyalty, so that from the four attributes, the results of the classification of customer satisfaction levels in the category of satisfied and dissatisfied can be obtained. By applying various equations and steps regarding the calculation of the C4.5 algorithm, namely by calculating entropy, split info, gain and gain ratio values with the attributes of Facility, Service, Loyalty, Category. The set of attributes, namely Very Satisfied, Satisfied, Quite Satisfied, Dissatisfied, Very Dissatisfied. Input data as much as 155 data. From this data, it is divided into 2, namely 100 training data and 55 testing data. customer satisfaction classification "Unsatisfied" as many as 91 customers, while the "Unsatisfied" as many as 9 customers.

Keywords: C4.5 Algorithm, Customer Satisfaction, Online Transportation, Data Mining

Abstrak- Klasifikasi kepuasan pelanggan pengguna transportasi online merupakan suatu sikap positif atau respon pelanggan terhadap pelayanan penggunaan transportasi online. Pengklasifikasian kepuasan pengguna transportasi online ini dilakukan dengan tujuan untuk meningkatkan kualitas perusahaan khususnya di bidang pelayanan pengguna transportasi roda dua. Dalam penelitian ini, tingkat kepuasan pelanggan diklasifikasikan dalam empat atribut yaitu harga, fasilitas, pelayanan, dan loyalitas sehingga dari ke empat atribut tersebut dapat diperoleh hasil pengklasifikasian tingkat kepuasan pelanggan dalam kategori puas dan tidak puas. Dengan menerapkan berbagai persamaan dan langkah-langkah mengenai perhitungan algoritma C4.5, yaitu dengan menghitung entropy, split info, gain dan nilai gain ratio dengan atribut Fasilitas, Pelayanan, Loyalitas, Kategori. Himpunan Atribut yaitu Sangat Puas, Puas, Cukup Puas, Tidak Puas, Sangat Tidak Puas. Input data sebanyak 155 data. Dari data tersebut maka dibagi 2 yaitu 100 data data latih (training) dan 55 data data uji (testing). klasifikasi kepuasan pelanggan "Tidak Puas" sebanyak 91 orang pelanggan, sedangkan yang "Tidak Puas" sebanyak 9 orang pelanggan.

Kata kunci: Algoritma C4.5, Kepuasan Pelanggan, Transportasi Online, Data Mining

1. PENDAHULUAN

Penyedia jasa transportasi online di Indonesia terus berkembang seperti *Maxim*, *FastGo*, *BitCar*, *Bonceng*, dan *Anterin.id*, *Gojek* dan *Grab*. Transportasi *online* merupakan pemindahan barang dan manusia dari tempat asal ke tempat tujuan. Salah satu jenis transportasi yang dimiliki *Maxim Transportation Online (Maxim)* adalah *GoBike*, memberikan layanan antar maupun jemput menggunakan roda dua. Untuk meningkatkan kualitas pelayanan terhadap pelanggan, maka perusahaan menyediakan fitur penilaian pada aplikasi transportasi *online*. Fitur penilaian berbentuk feedback, dapat dimanfaatkan pelanggan secara langsung untuk memberikan penilaian atau opini terhadap kualitas pelayanan yang diberikan oleh *Maxim Transportation Online*, salah satu perusahaan yang bergerak dalam bidang transportasi online yang berada di Kota Medan. *Maxim Transportation Online* mengalami beberapa hal kendala dan masalah dalam mengklasifikasikan tingkat kepuasan pelanggannya khususnya pelanggan pengguna roda dua (*GoBike*) dikarenakan tingkat kepuasan pelanggan yang kurang memuaskan salah satunya dalam bidang harga, fasilitas yang ditawarkan, pelayanan driver kepada pelanggan dan loyalitas perusahaan terhadap pelanggan [1].

Berdasarkan analisa tingkat kepuasan pelanggan, sekitar 17,53% pelanggan mengalami tidak puas atas penggunaan transportasi online periode tahun 2021 berdasarkan hasil angket yang disebarkan perusahaan kepada pelanggan pengguna *Maxim Transportation Online*. Contoh pengklasifikasian tingkat kepuasan pelanggan adalah dengan mengelompokkan satu per satu pelanggan yang puas dan tidak puas terhadap harga yang ditawarkan, fasilitas, pelayanan dan loyalitas driver terhadap pelanggan. Dari hasil klasifikasi tersebut kemudian dicatat dengan menggunakan aplikasi *Microsoft Office Excel* untuk melihat persentase kepuasan pelanggan dengan tujuan untuk memudahkan pelanggan pengguna transportasi online yang puas dan tidak puas. Untuk mengatasi permasalahan mengenai klasifikasi kepuasan pelanggan maka harus dibangun sebuah sistem aplikasi data mining dengan menggunakan Algoritma C4.5 agar bisa

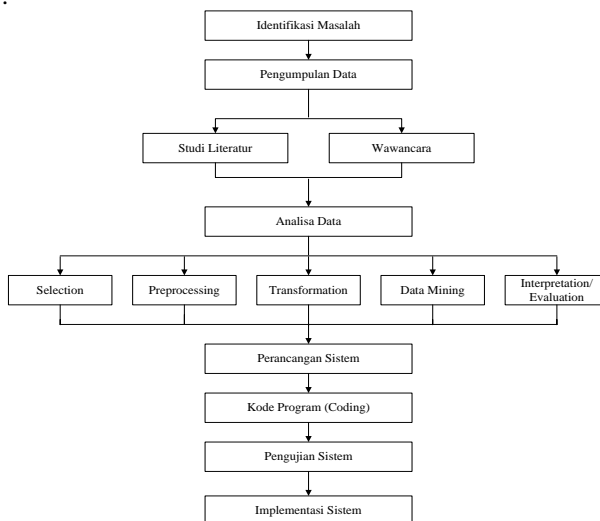
mengatasi dan mengelola data klasifikasi kepuasan pelanggan. Sistem ini nantinya dibuat untuk membantu mengklasifikasikan kepuasan pelanggan pengguna *GoBike* [2].

Data mining sebuah penelitian dari proses kumpulan data untuk mendapatkan hubungan yang tak diduga dan menyederhanakan data dengan cara yang berbeda dari sebelumnya, sehingga dapat dipahami dan memiliki manfaat bagi pemilik data. Data mining merupakan suatu cara mencari hubungan yang signifikan antara variabel dalam set data yang besar dengan memanfaatkan teknologi yang sesuai dan dalam penelitiannya memakai studi literatur yang berbeda. Beberapa teknik yang ada dalam data mining salah satunya yaitu *classification* [3]. Salah satu algoritma data mining yang digunakan untuk klasifikasi data adalah Algoritma *C4.5*. Algoritma *C4.5* merupakan salah satu contoh metode untuk membuat pohon keputusan dan tepat digunakan untuk mengklasifikasi data yang sangat besar ke dalam *class* tertentu menurut bentuk data yang ada. Data mining dan pembelajaran mesin (*machine learning*), algoritma *C4.5* digunakan untuk mempelajari data dengan jumlah yang besar dan menggunakan model pembelajaran berupa pohon keputusan [4]. Pohon keputusan adalah alternatif pemecahan masalah dalam beberapa teknik pada data mining salah satunya dalam bidang mengklasifikasikan tingkat kepuasan pelanggan [5].

Beberapa penelitian terdahulu yang telah berhasil menerapkan Algoritma *C4.5* dalam metode klasifikasi implementasi algoritma *C4.5* klasifikasi tingkat kepuasan pembeli *online shop*, menerapkan algoritma *C4.5* dalam memutuskan kepuasan pelanggan [6]. Jenis penelitian ini yakni klasifikasi dengan konsep data mining dengan melibatkan sebanyak 100 data yang kemudian diambil sampel sebanyak 15 data konsumen *polaroid vektor & gift* Jogja yang dikategorikan dengan: puas dan tidak puas. Hasil penelitian terdapat tingkat akurasi sebesar 91%, dengan nilai presisi pada prediksi puas sebesar 66.67% dan nilai presisi pada prediksi tidak puas sebesar 33.3%.

2. METODE PENELITIAN

Konsep pada penelitian yang saling berhubungan, penggambaran variabel satu dengan lainnya bisa terkoneksi secara detail dan sistematis, Gambar 1.



Gambar 1. Tahapan Penelitian

Uraian kerangka kerja penelitian :

1. Identifikasi Masalah

Identifikasi masalah adalah merupakan langkah pertama dalam melakukan suatu penelitian. Pada tahap identifikasi masalah harus benar-benar dirumuskan agar aplikasi penerapan data mining dengan menggunakan algoritma *C4.5* untuk mengklasifikasikan kepuasan pelanggan transportasi online benar-benar bermanfaat bagi *Maxim Transportation Online*, dengan melihat apa saja faktor atau kesulitan dalam mengklasifikasikan tingkat kepuasan pelanggan pengguna *GoBike* sehingga *Maxim Transportation Online* tidak mengalami kesulitan dan kesalahan data untuk menganalisa dan mengklasifikasikan tingkat kepuasan pelanggan transportasi *online* khususnya pelanggan pengguna *GoBike*.

2. Pengumpulan Data

a. Studi Literatur

Diperoleh dari buku, buku elektronik (*e-book*), jurnal ilmiah, artikel laporan penelitian, dan penelusuran situs-situs internet. *Output* dari studi literatur ini adalah terkoleksinya referensi yang relevan dengan rumusan masalah, agar memperkuat permasalahan serta sebagai dasar teori dalam melakukan studi dan juga menjadi dasar untuk melakukan desain kendali serta simulasi alur dalam aplikasi data mining.

b. Wawancara

Proses pengumpulan data dengan melakukan wawancara langsung pada bagian *Maxim Order* pada *Maxim Transportation Online*, adapun yang menjadi responden dalam proses wawancara ini yaitu Bapak Muhammad Farizi sebagai kepala *Maxim Order*.

3. Analisis Data

Setelah data selesai dikumpulkan maka dilakukan pengolahan data atau analisis data. Pada penelitian ini analisa data dilakukan dengan menggunakan Algoritma *C4.5*. Secara umum Algoritma *C4.5* memiliki beberapa langkah untuk membangun pohon keputusan:

- 1) Pilih atribut sebagai akar
- 2) Buat cabang untuk masing-masing nilai
- 3) Bagi kasus dalam cabang
- 4) Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama
- 5) Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus persamaan berikut ini:

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

- 6) Sebelum mendapatkan nilai *Gain* maka harus mencari nilai *Entropy* terlebih dahulu. *Entropy* digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar dari *Entropy* adalah sebagai berikut:

$$\text{Entropy}(S,A)= - \sum_{i=1}^n P_i * \text{Log}_2 P_i \tag{2}$$

Keterangan: S : Himpunan kasus A : Fitur n : Jumlah partisi P_i : Proporsi dari S_i terhadap S.

- 7) Apabila ada atribut yang mempunyai banyak nilai atribut perlu untuk menghitung *gain ratio*, sebelumnya perlu kita ketahui suatu istilah baru yang disebut *split information*, yang bisa dihitung dengan rumus persamaan sebagai berikut :

$$\text{Split Info}(S,A)= \sum_{i=1}^c \frac{S_i}{S} \text{Log}_2 \frac{S_i}{S} \tag{3}$$

Keterangan : S = ruang (data) sampel yang digunakan untuk training A = atribut S_i = jumlah sampel untuk atribut i, S_i sampai S_c adalah subset c yang dihasilkan dari pemecahan S dengan menggunakan atribut A yang mempunyai sebanyak c nilai. Selanjutnya *gain ratio* dihitung dengan cara:

$$\text{Gain Ratio}(S,A)= \frac{\text{Gain}(S,A)}{\text{SplitInfo}(S,A)} \tag{4}$$

Langkah-langkah analisa data dengan Algoritma *C4.5* :

- 1) *Data selection*
Seleksi (pemilihan) data dari data yang telah diperoleh. Data yang akan diseleksi merupakan data tingkat kepuasan pelanggan pengguna transportasi *online* khususnya *GoBike* dari *Maxim Transportation Online*. Data tersebut diseleksi atribut dengan memilih atribut yang sesuai pada kebutuhan penelitian dengan tujuan untuk memperkecil ruang lingkup penelitian [7]
- 2) *Pre-Processing*
Sebelum proses data mining dilaksanakan, perlu dilakukan proses cleaning atau proses pembersihan data yaitu pembersihan data tingkat kepuasan pelanggan pengguna transportasi online yang telah diperoleh sebelumnya. Proses pembersihan data mencakup dimana informasi yang tidak dibutuhkan dibuang, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data [8].
- 3) *Transformation*
Setelah data dipilih maka akan dilakukan proses transformasi data yaitu dilakukan perubahan data yang memiliki tipe data yang awalnya tidak bisa diolah secara matematis menjadi data yang bisa diolah dengan tujuan untuk menghindari data yang rusak dan tidak valid dalam melakukan klasifikasi tingkat kepuasan pelanggan [9].
- 4) *Data Mining*
Data mining merupakan proses mencari pola atau informasi yaitu menarik data terpilih dengan menggunakan teknik atau metode tertentu. Digunakan Algoritma *C4.5* untuk melakukan klasifikasi kepuasan pelanggan transportasi *online*. Tujuan Algoritma *C4.5* adalah untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru [10].
- 5) *Interpretation/Evaluation*
Implementasi atau evaluasi pola yang dihasilkan dari proses data mining [11]. Algoritma *C4.5* perlu ditampilkan dalam bentuk yang mudah dipahami oleh pihak yang akan menggunakan sistem yang dibuat yaitu pihak *Maxim Order* pada *Maxim Transportation Online*.
4. Perancangan Sistem
Perancangan sistem merupakan sekumpulan aktivitas yang menggambarkan secara rinci bagaimana sistem akan berjalan. Sistem yang dirancang dengan menggunakan pemodelan UML (*Unified Modeling Language*) yang meliputi *use case* yang merupakan gambaran fungsionalitas, Activity diagram yang merupakan alur dari sistem yang

dibuat, class diagram sebagai gambaran objek yang mewakili sistem yang digunakan sebagai dasar membangun sistem.

5. Kode Program (*Coding*)
 Tahap pembuatan *coding* program. Pembuatan *coding* dimaksudkan untuk menyederhanakan judul kolom dalam proses entry data (memasukkan atau tabulasi data). Dalam pengkodean program menggunakan bahasa pemrograman PHP dengan database *MySQL*.
6. Pengujian Sistem
 Pengujian sistem adalah menguji fungsionalitas dari semua sistem dan penanganan kesalahan pada sistem yang dibuat apakah sudah sesuai dengan yang diharapkan. Pengujian sistem ini nantinya pertama-tama dilakukan pada server localhost, misalnya :*http://localhost/klasifikasi-kepuasan-pelanggan/*
7. Implementasi Sistem
 Implementasi merupakan langkah akhir untuk dilakukan dalam proses penerapan sistem baru dimana tahap ini merupakan tahap meletakkan sistem supaya siap untuk dioperasikan dan dapat dipandang sebagai usaha untuk mewujudkan sistem yang telah dirancang.

3. HASIL DAN PEMBAHASAN

Data diperoleh dari bagian *Maxim Order* sebagai input data yang akan diproses dengan jumlah sebanyak 155 data. Dari data tersebut maka dibagi 2 yaitu 100 data untuk data latih (training) dan 55 data untuk data uji (testing). Skala *Likert* Kepuasan Pelanggan, Tabel 1.

Tabel 1. Skala *Likert* Kepuasan Pelanggan

Pilihan	Star	Singkatan	Skor
Sangat Puas	★★★★★	SP	5
Puas	★★★★	P	4
Cukup Puas	★★★	CP	3
Tidak Puas	★★	TP	2
Sangat Tidak	★	STP	1

Tahap seleksi data dilakukan untuk mendapatkan data yang sesuai dengan kebutuhan dalam mengklasifikasikan kepuasan pelanggan pengguna transportasi online, dibutuhkan data hasil klasifikasi kepuasan pelanggan yang di dapat dari pengumpulan data sebelumnya. Dalam tahap seleksi data ini terlebih dahulu ditentukan jenis-jenis atribut klasifikasi dengan tujuan agar sesuai kebutuhan sistem yang akan dibangun, Harga (A1), Fasilitas (A2), Pelayanan (A3), Loyalitas (A4), Kategori (A5). Setelah ditentukan atribut dalam pengklasifikasian kepuasan pelanggan pengguna transportasi online, selanjutnya mengelompokkan atribut, Tabel 2.

Tabel 2. Nilai Atribut

No	Nama Atribut	Himpunan Atribut	Nilai Atribut
1	Harga	Sangat Puas	5
		Puas	4
		Cukup Puas	3
		Tidak Puas	2
		Sangat Tidak Puas	1
2	Fasilitas	Sangat Puas	5
		Puas	4
		Cukup Puas	3
		Tidak Puas	2
		Sangat Tidak Puas	1
3	Pelayanan	Sangat Puas	5
		Puas	4
		Cukup Puas	3
		Tidak Puas	2
		Sangat Tidak Puas	1
4	Loyalitas	Sangat Puas	5
		Puas	4

		Cukup Puas	3
		Tidak Puas	2
		Sangat Tidak Puas	1
5	Kategori	Puas	
		Tidak Puas	

Pada data mining terdapat data yang tidak lengkap. Data yang tidak lengkap disebabkan karena ada data yang kosong atau atribut yang salah. Data pengklasifikasian kepuasan pelanggan pengguna transportasi online pada Maxim Transportation Online, ada sebagian atribut yang tidak perlu sehingga proses data preprocessing perlu dilakukan dengan tujuan agar database sesuai dengan ketentuan. Algoritma C4.5, pohon keputusan (*Decision tree*) memiliki dua jenis atribut dari data yang terdiri dari beberapa atribut input dan atribut target, fungsinya sebagai pembandingan dalam perhitungan Gain dan Ratio, Tabel 3.

Tabel 3. Pengelompokkan Jumlah Kasus Setiap Atribut Klasifikasi

Atribut	Jumlah Kasus	Puas	Tidak Puas
Total	100	74	26
Harga	1	4	0
	2	6	2
	3	23	18
	4	49	40
	5	18	14
Fasilitas	1	3	2
	2	5	4
	3	21	16
	4	20	14
	5	51	38
Pelayanan	1	3	2
	2	7	6
	3	30	23
	4	24	17
	5	36	26
Loyalitas	1	6	5
	2	10	9
	3	17	12
	4	22	16
	5	45	32

Langkah-langkah dalam menentukan node akar tersebut adalah sebagai berikut:

a. Perhitungan *Entropy*

Langkah awal algoritma C4.5 adalah mencari nilai *entropy*, menggunakan rumus persamaan :

$$Entropy(S) = \left(-\frac{S_1}{S} * \text{Log}_2 \left(\frac{S_1}{S}\right)\right) + \left(-\frac{S_2}{S} * \text{Log}_2 \left(\frac{S_2}{S}\right)\right) \tag{5}$$

b. Perhitungan *Information Gain*

Hasil perhitungan *information gain* Information Gain Atribut Harga = 0.123, Information Gain Atribut Fasilitas = 0.003, Information Gain Atribut Pelayanan = 0.007, Information Gain Atribut Loyalitas = 0.016.

c. Perhitungan *Split Information*

Split information digunakan sebagai pembagi dari *Gain (A)* yang akan menghasilkan *gain ratio*. perhitungan *split info* dari setiap atribut dengan menggunakan rumus persamaan berikut ini:

$$Split\ Info = - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Log}_2 \frac{|S_i|}{|S|} \tag{6}$$

Hasil perhitungan Split Info Atribut Harga = 1.867, Split Info Atribut Fasilitas = 1.801, Split Info Atribut Pelayanan = 1.966, Split Info Atribut Loyalitas = 2.009.

d. Perhitungan *Gain Ratio*

Perhitungan *gain ratio* dalam mengklasifikasi kepuasan pelanggan transportasi online :

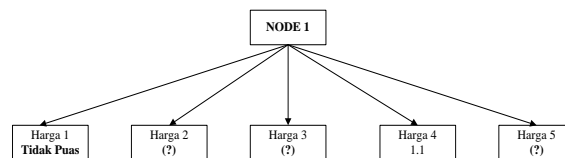
$$GainRatio (A) = \frac{Gain(A)}{SplitInfo(A)} \tag{7}$$

Hasil perhitungan GainRatio Atribut Harga = 0.066, GainRatio Atribut Fasilitas = 0.002, GainRatio Atribut Pelayanan = 0.003, GainRatio Atribut Loyalitas = 0.008. Berdasarkan perhitungan nilai *entropy*, *information gain*, *split info*, dan *gain ratio* diatas, maka dapat ditampilkan dalam bentuk Tabel 4.

Tabel 4. Hasil Perhitungan Node 1 Akar

NODE	Nilai Atribut	Total Kasus	Puas (S1)	Tidak Puas (S2)	Entropy	Gain	Split Info	Gain Ratio
1	TOTAL	100	74	26	0.82			
Harga	1	4	0	4	0	0.123	1.867	0.066
	2	6	2	4	0.91			
	3	23	18	5	0.75			
	4	18	14	4	0.76			
	5	49	40	9	0.68			
Fasilitas	1	3	2	1	0.91	0.003	1.801	0.002
	2	5	4	1	0.72			
	3	21	16	5	0.79			
	4	20	14	6	0.88			
	5	51	38	13	0.81			
Pelayanan	1	3	2	1	0.91	0.007	1.966	0.003
	2	7	6	1	0.59			
	3	30	23	7	0.78			
	4	24	17	7	0.87			
	5	36	26	10	0.85			
Loyalitas	1	6	5	1	0.65	0.016	2.009	0.008
	2	10	9	1	0.46			
	3	17	12	5	0.87			
	4	22	16	6	0.84			
	5	45	32	13	0.86			

Diketahui bahwa nilai *gain ratio* tertinggi berada pada atribut “Harga” dengan nilai sebesar 0.066, sehingga atribut harga menjadi *node* akar selanjutnya. Pada atribut “Harga” terdapat lima nilai atribut yaitu 1, 2, 3, 4 dan 5. Nilai atribut “4” sudah mengklasifikasikan kasus menjadi satu yaitu keputusannya “Puas” atau “Tidak Puas” sehingga didapatkan pohon keputusan, Gambar 2.



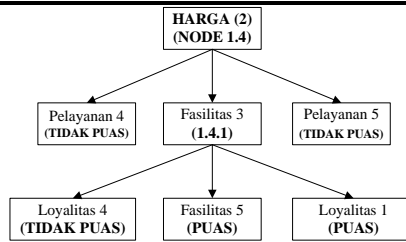
Gambar 1. Pohon Keputusan Node 1

Berdasarkan perhitungan nilai *entropy*, *information gain*, *split information* dan nilai *gain ratio* maka dapat dilihat pada Tabel 5.

Tabel 5. Hasil Perhitungan Node 1.4.1

NODE	Nilai Atribut	Total Kasus	Puas (S1)	Tidak Puas (S2)	Entropy	Gain	Split Info	Gain Ratio
1.4.1 PELAYANAN (3)	TOTAL	3	2	1	0.918			
Fasilitas	3	1	1	0	0	0.252	0.918	0.274
	5	2	1	1	1			
Loyalitas	1	1	1	0	0	0.918	1.585	0.579
	4	1	0	1	0			
	5	1	1	0	0			

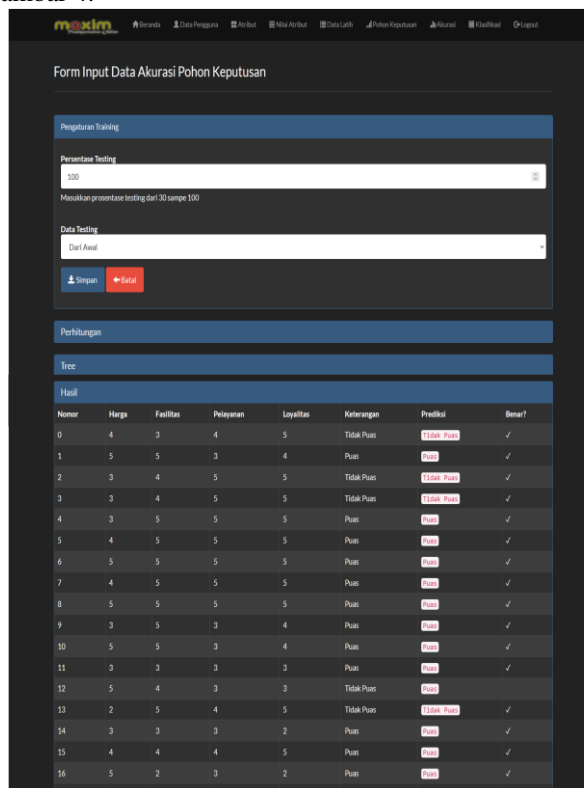
Hasil perhitungan *node* 1.4.1 diketahui bahwa atribut yang memiliki nilai *gain ratio* tertinggi yaitu atribut “Pelayanan” dengan nilai sebesar 0.06 sehingga dengan demikian maka atribut “Pelayanan” dijadikan sebagai *node* akar berikutnya, Gambar 3.



Gambar 3. Pohon Keputusan Node 1.4.1

Berdasarkan tabel hasil klasifikasi kepuasan pelanggan pengguna transportasi online diatas, maka dapat disimpulkan bahwa jumlah pelanggan pengguna transportasi online dengan klasifikasi “Tidak Puas” sebanyak 91 orang pelanggan, sedangkan yang “Tidak Puas” sebanyak 9 orang pelanggan.

Implementasi sistem, program di hosting ke server agar bisa di akses. Misalnya pada server localhost: <http://localhost/klasifikasi-kepuasan-pelanggan-maxim/>. Implementasi Output Hasil Uji Akurasi Pohon Keputusan merupakan tampilan implementasi form output hasil uji akurasi pengklasifikasian kepuasan pelanggan pengguna transportasi online, Gambar 4.



Gambar 4. Implementasi Output Hasil Uji Akurasi Pohon Keputusan

Aplikasi data mining untuk mengklasifikasikan kepuasan pelanggan pengguna transportasi online telah berhasil diimplementasikan pada sistem dengan menerapkan Algoritma C4.5 hasil klasifikasi kepuasan pelanggan pengguna transportasi online, jumlah pelanggan pengguna transportasi online dengan klasifikasi “Tidak Puas” sebanyak 91 orang pelanggan, sedangkan yang “Tidak Puas” sebanyak 9 orang pelanggan.

4. KESIMPULAN

Kesimpulan penelitian :

1. Aplikasi data mining untuk mengklasifikasikan kepuasan pelanggan pengguna transportasi online dilakukan dengan menggunakan Algoritma C4.5 yaitu dengan menghitung nilai entropy, split info, gain dan gain ratio dengan hasil yang diperoleh berupa pohon keputusan.
2. Aplikasi data mining klasifikasi kepuasan pelanggan pengguna transportasi online ini dapat menyediakan informasi hasil analisa dan hasil perhitungan metode Algoritma C4.5 serta informasi pohon keputusan bagi admin dengan pengujian terhadap 100 (seratus) data latih dengan empat atribut klasifikasi yaitu harga, fasilitas, pelayanan dan

loyalitas, maka diperoleh hasil pengklasifikasian kepuasan pelanggan pengguna transportasi online dalam bentuk pohon keputusan.

REFERENSI

- [1] A. Febriyani, G. K. Prayoga, and ..., "Index Kepuasan Pelanggan Informa dengan Menggunakan Algoritma C. 45," JURIKOM (Jurnal ..., vol. 8, no. 6, pp. 330–335, 2021, doi: 10.30865/jurikom.v8i6.3686.
- [2] A. S. Sinaga, M. H. Munandar, and A. S. Sitio, "Machine learning algorithm to identifies fraud emails with feature selection," IOP Conf. Ser. Mater. Sci. Eng., vol. 1088, no. 1, p. 012011, 2021, doi: 10.1088/1757-899x/1088/1/012011.
- [3] Asmaul Husnah Nasrullah, "penerapan metode C45 untuk klasifikasi mahasiswa drop out," Ilk. J. Ilm., vol. 10, pp. 244–250, 2018.
- [4] E. Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti," J. Online Inform., vol. 2, no. 1, p. 36, 2017, doi: 10.15575/join.v2i1.71.
- [5] H. S. P. H. O. R. K. Sormin, "Implementasi Algoritma Decision Tree Dalam Menentukan Lokasi Penjualan Dagangan Pada Dinas Pasar Serbelawan Simalungun," SmartEDU Yayasan Adwitiya Basurata Inov., vol. 1, no. 2, pp. 61–69, 2022.
- [6] M. S. Anita Sindar RM Sinaga, "Machine Learning Prediksi Karakter Pengguna Hastag (#) Bahasa Generasi Milenial Di Sosial Media," News.Ge, p. <https://news.ge/anakliis-porti-aris-qveynis-momava>, 20189.
- [7] S. A. A. Sherly Maisa Putri and Fakultas, "Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus: Hinet Batam)."
- [8] S. Al Syahdan and A. Sindar, "Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota," J. Nas. Komputasi dan Teknol. Inf., vol. 1, no. 2, 2018, doi: 10.32672/jnkti.v1i2.771.
- [9] S. Dutalia, A. K. Lalo, P. Batarius, Y. Carmeneja, and H. Siki, "Implementasi Algoritma C4 . 5 Untuk Klasifikasi Penjualan," vol. 06, pp. 1–12, 2021.
- [10] S. Takalapeta, "Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5," J I M P - J. Inform. Merdeka Pasuruan, vol. 3, no. 3, pp. 34–38, 2018, doi: 10.37438/jimp.v3i3.186.
- [11] Y. Bastian, H. S. Tambunan, and W. Saputra, "Analisis Penerapan Algoritma C4 . 5 Dalam Mengukur Tingkat Kepuasan Pelanggan Indihome Pada Kota Pematangsiantar," KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen), vol. 2, no. 1, pp. 62–69, 2021.