

Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus: Hinet Batam)

Sherly Maisa Putri, Sasa Ani Arnomo

Fakultas Teknik, Program Studi Sistem Informasi, Universitas Putera Batam, Batam, Indonesia

Email: Sherlymaisaa6596@gmail.com

Abstrak—Kualitas pelayanan merupakan tujuan dari perusahaan dalam memberikan kepuasan kepada konsumen baik perusahaan jasa maupun non jasa. Salah satu perusahaan jasa penyedia layanan akses internet yang berkomitmen untuk meningkatkan kualitas pelayanan terhadap kepuasan konsumen adalah Hinet Batam. Penelitian ini bertujuan untuk menganalisa kualitas pelayanan terhadap kepuasan konsumen Hinet Batam dengan menggunakan algoritma C4.5. Variabel penilaian mencakup empathy, responsiveness, reliability dengan keputusan puas dan tidak puas. Hasil dari pengolahan metode Data Mining dengan menggunakan algoritma C4.5 ternyata mampu mengklasifikasi tingkat kepuasan konsumen terhadap kualitas Pelayanan. Dengan nilai presisi prediksi puas sebesar 61.7% dan nilai presisi prediksi tidakpuas 38,3%. Hasil penelitian ini bisa dijadikan acuan untuk meningkatkan kualitas pelayanan serta untuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan.

Kata Kunci: Kualitas Pelayanan, Kepuasan Konsumen, Data Mining, Algoritma C4.5, Pohon Keputusan.

Abstract—Quality of service is the goal of the company in providing satisfaction to consumers both service and non-service companies. One of the internet access service provider companies that is committed to improving service quality to customer satisfaction is Hinet Batam. This study aims to analyze the quality of service to Hinet Batam consumer satisfaction using the C4.5 algorithm. Valuation variables include empathy, responsiveness, reliability with satisfied and dissatisfied decisions. The results of processing the Data Mining method using the C4.5 algorithm were able to classify the level of customer satisfaction with service quality. With a satisfied predictive precision value of 61.7% and an unsatisfied predictive precision value of 38.3%. The results of this study can be used as a reference to improve service quality and to maintain customer loyalty in using internet service data package products and increase sales.

Keywords: Quality of Service, Customer Satisfaction, Data Mining, Algorihtm C4.5, Descision Tree.

1. PENDAHULUAN

Perkembangan teknologi telekomunikasi dan informasi saat ini sangat berkembang dengan pesat salah satunya yaitu Internet. Terbukti dengan banyaknya kehadiran perusahaan yang menyediakan layanan internet atau sering disebut dengan *Internet Service Provider (ISP)*. Salah satu *provider* yang ada Indonesia yaitu PT. Berca Hardayaperkasa dengan produk layanan akses data internet yaitu Hinet. Hinet telah berkembang di beberapa wilayah di daerah Indonesia salah satunya yaitu kota Batam. Hinet Batam berdiri sejak Januari 2019 dan sudah memiliki 4 *Base Traineceiver Station (BTS)* yang tersebar di 4 kecamatan kota Batam.

Dengan banyaknya perusahaan penyedia jasa layanan di kota Batam yang sudah mendapatkan *branding* / memiliki *image* bagus bagi masyarakat kota Batam membuat Hinet Batam harus berusaha keras untuk mempertahankan loyalitas konsumen bahkan meningkatkan penjualan.

Agar mampu mempertahankan loyalitas konsumen bahkan meningkatkan penjualan, Hinet Batam harus mampu untuk meningkatkan kepuasan konsumen. Kepuasan konsumen mencakup perbedaan antara tingkat kepentingan kinerja atau hasil yang dirasakan, serta merupakan evaluasi purna beli dimana alternatif yang dipilih sekurang-kurangnya dapat memberikan hasil sama atau melampaui harapan. Sedangkan ketidakpuasan dapat terjadi apabila hasil yang diperoleh tidak memenuhi harapan yang diinginkan konsumen [1].

Untuk menciptakan kepuasan konsumen diperlukan beberapa faktor, salah satunya adalah kualitas pelayanan. Kualitas pelayanan adalah tingkat mutu yang diharapkan dan pengendalian keragaman dalam mencapai mutu tersebut untuk memenuhi kebutuhan konsumen. Kualitas pelayanan yang melebihi harapan konsumen akan membuat konsumen merasa puas [2].

Masalah pelayanan sebenarnya bukanlah hal yang sulit atau rumit, tetapi apabila hal ini tidak diperhatikan maka dapat menimbulkan permasalahan bagi perusahaan. Untuk itu perlu adanya sebuah riset kepuasan konsumen sebagai evaluasi bagi perusahaan.

Berdasarkan masalah di atas prediksi kualitas pelayanan terhadap kepuasan konsumen dapat diproses melalui penerapan teknologi *data mining* dengan algoritma C4.5. *Data Mining* merupakan penambangan atau penemuan informasi baru dengan mencari pola antar data secara otomatis. *Data Mining* memiliki beberapa teknik salah satunya yaitu klasifikasi. Teknik klasifikasi terdiri dari beberapa metode, dan *decision tree* adalah salah satu dari metode klasifikasi. Kemudian metode *decision tree* memiliki beberapa algoritma salah satunya yaitu algoritma C4.5 [3].

2. METODOLOGI PENELITIAN

2.1 Pengumpulan Data

Penelitian menggunakan beberapa metode pendukung guna mendapatkan data pada objek penelitian. Metode yang digunakan adalah:

- a. Wawancara
Wawancara adalah teknik pengumpulan data yang dilakukan dengan menemui dan tanya jawab langsung antara peneliti dan narasumber. Pada metode ini peneliti langsung menanyakan beberapa hal tentang keluhan dari konsumen.
- b. Kuesioner
Kuesioner merupakan teknik pengumpulan data dengan seperangkat pertanyaan yang akan diisi oleh responden. Kuesioner pada penelitian ini disebarkan kepada responden yang merupakan konsumen Hinet Batam secara online dengan menggunakan aplikasi dari Google yaitu *Google Form*.
- c. Studi Pustaka
Studi pustaka merupakan kegiatan untuk mencari informasi yang relevan dengan topik atau masalah yang menjadi objek penelitian. Informasi tersebut dapat diperoleh dari buku-buku, karya ilmiah dan sumber-sumber lain yang berkaitan dengan kualitas pelayanan, kepuasan konsumen, *data mining*, dan algoritma C4.5.

2.2 Data Mining

Menurut Turban dalam bukunya yang berjudul "*Decision Support Systems and Intelligent Systems*", *data mining* adalah suatu hasil yang digunakan untuk menguraikan penemuan pengetahuan di dalam *basis data*. *Data mining* adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *basis data* besar. Dapat disimpulkan *data mining* merupakan suatu proses otomatis untuk menganalisa data. Data yang akan diproses merupakan data yang sangat besar, sehingga sulit untuk memproses dengan manual [6].

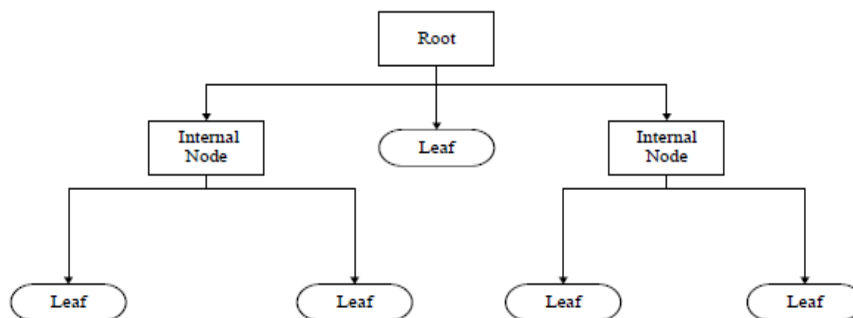
Data mining, sering juga disebut sebagai *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam *set data* berukuran besar. *Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi [7].

2.2.1 Klasifikasi

Klasifikasi merupakan proses untuk menyatakan suatu objek ke salah satu kategori yang sudah didefinisikan sebelumnya. Klasifikasi juga bisa didefinisikan sebagai proses pembelajaran fungsi target (model klasifikasi yang memetakan setiap sekumpulan atribut x (*input*) ke salah satu kelas yang didefinisikan sebelumnya). *Input* didefinisikan sebagai sekumpulan *record* (*training set*), dan setiap *record* terdiri atas sekumpulan atribut adalah kelas [3].

2.2.2 Pohon Keputusan

Pohon keputusan dengan algoritma klasifikasi sangat populer dan berpengaruh dibidang sistem informasi. Pohon keputusan digunakan untuk membangun pohon prediksi suatu analisis[8]. Pohon keputusan (*Decision tree*) adalah *flowchart* seperti struktur *tree*, dimana tiap *internal node* menunjukkan sebuah *test* pada sebuah atribut, tiap cabang menunjukkan hasil *test*, dan *leaf node* menunjukkan *class-class* atau *class distribution* [1].



Gambar 1. Model Pohon Keputusan

2.3 Kualitas Pelayanan

Terdapat 5 dimensi pokok atau factor-faktor yang mempengaruhi kualitas pelayanan[4], diantaranya:

- a. Keandalan (*reliability*),
- b. Daya tanggap (*responsiveness*),
- c. Jaminan (*assurance*),
- d. Kemudahan akses (*emphaty*),
- e. Penampilan fisik (*tangible*).

2.4 Kepuasan Konsumen

Kepuasan konsumen adalah hasil yang dirasakan oleh pembeli yang mengalami kinerja sebuah perusahaan yang sesuai dengan harapannya. Pelanggan merasa puas jika harapan mereka dapat terpenuhi. Secara sederhana kepuasan konsumen dapat diartikan sebagai upaya pemenuhan sesuatu atau membuat sesuatu memadai [5].

2.5 Algoritma C4.5

Algoritmas C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedangkan pohon keputusan diartikan sebagai suatu cara untuk mempredikasi atau mengelompokkan yang sangat kuat [9]. Rumus algoritma C4.5 terbagi menjadi 2 rumus yaitu untuk mencari nilai *gain* dan untuk mencari nilai *entropy* [3].

Rumus untuk mencari nilai *gain* sebagai berikut:

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

Rumus untuk mencari nilai *entropy* sebagai berikut:

$$Entropy(s) = \sum_{i=1}^n - p_i * Log_2 p_i \tag{2}$$

3. HASIL DAN PEMBAHASAN

3.1 Analisis Data

Data kepuasan konsumen memiliki format sebagai berikut:

1. Pelayanan 1, merupakan kemudahan menghubungi layanan konsumen (*emphaty*)
2. Pelayanan 2, merupakan respons karyawan saat menghubungi layanan konsumen (*responsiveness*)
3. Pelayanan 3, merupakan respons karyawan saat memperbaiki kerusakan (*reliability*)
4. Hasil: Puas dan Tidak Puas

Tabel 1. Data Setelah Pra-Proses/Format Akhir

Pelayanan1	Pelayanan2	Pelayanan3	Hasil
Baik	Baik	Baik	Puas
Baik	Sangat Baik	Sangat Baik	Tidak Puas
Baik	Baik	Sangat Baik	Puas
Baik	Cukup	Baik	Tidak Puas
Cukup	Cukup	Cukup	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Baik	Cukup	Baik	Tidak Puas
Baik	Baik	Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Baik	Baik	Puas
Cukup	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Baik	Baik	Tidak Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Kurang	Baik	Tidak Puas
Cukup	Baik	Cukup	Puas
Kurang	Kurang	Kurang	Tidak Puas
Baik	Baik	Baik	Puas
Cukup	Baik	Baik	Puas
Cukup	Baik	Baik	Puas
Cukup	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Cukup	Baik	Tidak Puas
Cukup	Cukup	Cukup	Tidak Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Baik	Baik	Baik	Tidak Puas

Baik	Baik	Baik	Puas
Baik	Sangat Baik	Sangat Baik	Puas
Kurang	Kurang	Kurang	Tidak Puas
Cukup	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Cukup	Cukup	Tidak Puas
Cukup	Cukup	Cukup	Tidak Puas
Cukup	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Kurang	Cukup	Cukup	Tidak Puas
Kurang	Kurang	Kurang	Tidak Puas
Kurang	Kurang	Kurang	Tidak Puas
Cukup	Cukup	Cukup	Tidak Puas
Cukup	Cukup	Cukup	Tidak Puas
Baik	Baik	Baik	Puas
Kurang	Kurang	Kurang	Tidak Puas

Setelah dianalisis, data set tersebut memiliki 47 kasus yang terdiri dari 29 “PUAS” dan 18 “TIDAK PUAS” yang terdapat pada kolom hasil. Langkah selanjutnya yang harus dilakukan adalah menghitung entropinya.

Tabel 2. Tabel Data Set

Total Data	Puas	Tidak Puas	Entropy	Gain
47	29	18		

$$\begin{aligned}
 Entropy (S) &= - \frac{Entropy\ total\ (s)=29}{47} \times (2 \log 29 - 2 \log 47) + (-) \frac{18}{47} \times (2 \log 18 - 2 \log 47) \\
 &= (- 0.617021277 \times - 0.696607857) + (- 0.382978723 \times - 1.38466385) \\
 &= 0.429821869 + 0.530296794 \\
 &= \mathbf{0.960118663}
 \end{aligned}$$

Tabel 3. Hasil Perhitungan Entropy

Setelah	Total Data	Puas	Tidak Puas	Entropy	Gain
	47	29	18	0.960118663	

entropy total dihitung, maka langkah selanjutnya adalah melakukan analisis dan perhitungan entropy setiap atribut dan nilai-nilainya. Setelah entropy setiap atribut dihitung langkah selanjutnya menghitung gain pada setiap atribut menggunakan rumus berikut ini:

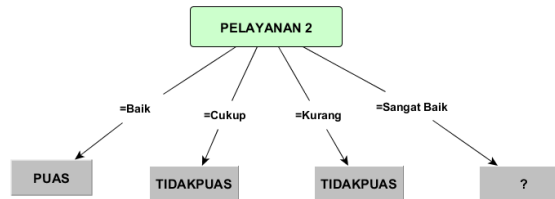
$$\begin{aligned}
 Gain\ Pelayanan\ 1 &= 0.960118663 - ((\frac{2}{47} \times 0) + ((\frac{20}{47}) \times 0.721928095) + ((\frac{17}{47}) \times 0.997502546) + ((\frac{6}{47}) \\
 &\times 0) + ((\frac{0}{47}) \times 0)) \\
 &= \mathbf{0.292116425}
 \end{aligned}$$

Tabel 4. Hasil Perhitungan Total

Atribut	Nilai	Total	Puas	Tidak Puas	Entropy	Gain
Pelayanan 1	Sangat Baik	2	2	0	0	0.292116425
	Baik	20	16	4	0.721928095	
	Cukup	17	9	8	0.997502546	
	Kurang	6	0	6	0	
Pelayanan 2	Sangat Kurang	0	0	0	0	0.618424881
	Sangat Baik	6	5	1	0.219195338	
	Baik	25	23	2	0.40217919	
	Cukup	10	1	9	0.468995594	
Pelayanan 3	Kurang	6	0	6	0	0.294896192
	Sangat Kurang	0	0	0	0	
	Sangat Baik	7	6	1	0.591672779	
	Baik	27	21	6	0.764204507	

Cukup	8	2	6	0.811278124
Kurang	5	0	5	0
Sangat Kurang	0	0	0	0

Setelah nilai *gain* pada setiap atribut diketahui maka langkah selanjutnya adalah membuat pohon keputusannya. Untuk membuat pohon keputusan tersebut langkah awalnya kita harus mencari nilai *gain* yang terbesar. Karena nilai *gain* Pelayanan2 yang paling besar maka pelayanan2 menjadi *node* akar (*root node*).



Gambar 2. Pohon Keputusan Node 1(Root Node)

Berdasarkan hasil dari *node* 1(*root node*), maka langkah selanjutnya yaitu menganalisis *node* 1.2. Untuk mempermudah analisisnya, tabel 1 difilter dengan data yang memiliki Pelayanan 2 = Sangat Baik Sehingga terbentuklah tabel berikut.

Tabel 5. Penyederhanaan Atribut Pelayanan 2

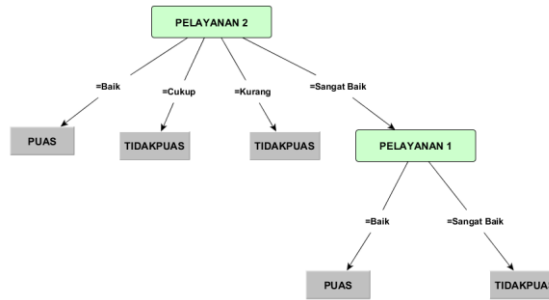
Pelayanan 2	Pelayanan 1	Pelayanan3	Hasil
Sangat Baik	Baik	Sangat Baik	Tidak Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Puas
Sangat Baik	Baik	Sangat Baik	Puas

Selanjutnya, data pada tabel 5 dianalisis dan dihitung kembali *entropy* dan nilai *gain*nya pada setiap atribut lainnya dengan mengecualikan atribut pelayanan 2 yang sudah berada pada jalur di atasnya sehingga membentuk tabel baru seperti di bawah ini.

Tabel 4. Hasil Perhitungan Node 1.2

Pelayanan2 Sangat Baik	Puas	Tidak Puas	Entropy			
6	5	1	0.650022422			
Atribut	Nilai	Total	Puas	Tidak Puas	Entropy	Gain
Pelayanan 1	Sangat Baik	4	4	0	0	0.316689088
	Baik	2	1	1	1	
	Baik	0	0	0	0	
	Cukup	0	0	0	0	
	Kurang	0	0	0	0	
Node 1.2 Pelayanan 3	Sangat Kurang	6	5	1	0.650022422	0
	Baik	0	0	0	0	
	Baik	0	0	0	0	
	Cukup	0	0	0	0	
	Kurang	0	0	0	0	

Dari tabel hasil perhitungan tersebut dapat diketahui bahwa *gain* Pelayanan 1 yang paling besar dan memiliki nilai **0.316689088**. Adapun pohon keputusan yang terbentuk dari analisis *node* 1.2 seperti gambar berikut.



Gambar 3. Pohon Keputusan Node 2

Dari hasil perhitungan *node* 1.2 tidak adalagi atribut yang bisa dianalisis karena atribut akhir sudah bernilai 0.

3.2 Rule Hasil Dari Perhitungan Entropy dan Gain

Adapun *rule* yang terbentuk berdasarkan pohon keputusan terakhir adalah sebagai berikut:

1. Jika pelayanan respons karyawan baik maka konsumen merasa puas
2. Jika pelayanan respons karyawan cukup maka konsumen tidak puas
3. Jika pelayanan respons karyawan kurang maka konsumen tidak puas
4. Jika pelayanan respons karyawan sangat baik dan kemudahan menghubungi layanan sangat baik maka konsumen puas
5. Jika pelayanan respons karyawan sangat baik dan kemudahan menghubungi layanan baik maka konsumen tidak puas

Untuk menghitung presentase PUAS dan TIDAK PUAS menggunakan rumus:

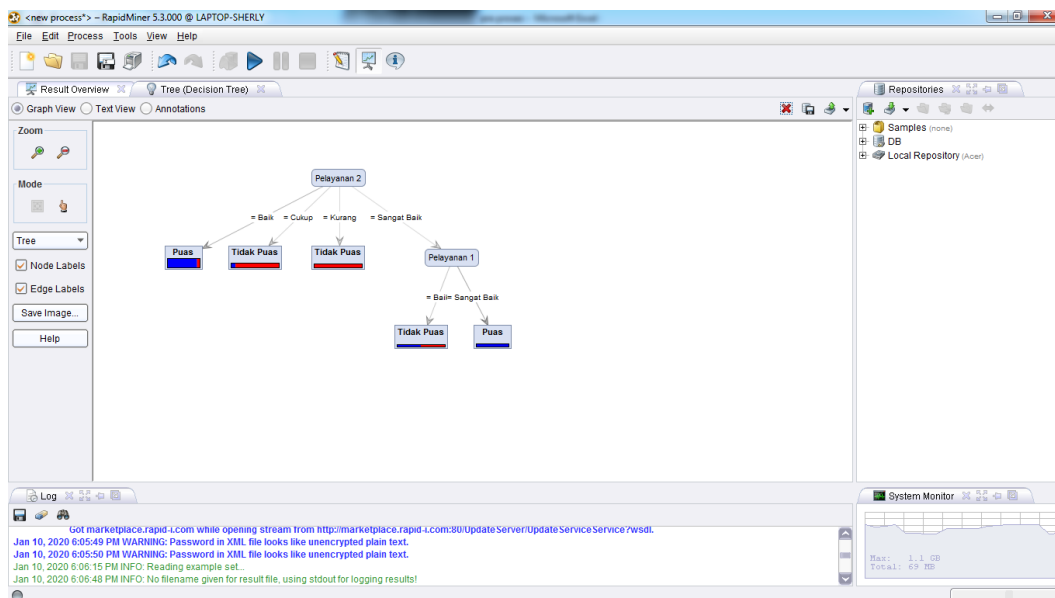
$$\text{Puas} = \frac{29}{47} \times 100 \% = 61.7\% \quad (3)$$

$$\text{Tidak Puas} = \frac{18}{47} \times 100 \% = 38.3\% \quad (4)$$

3.3 Pengujian Menggunakan Software RapidMiner

Pengujian terhadap analisa sangat penting karena apakah pengujian tersebut sesuai dengan yang diharapkan. Untuk menguji kebenaran hasil pengolahan data manual, maka dapat menggunakan *software RapidMiner* 5.3.000.

Berikut adalah pohon keputusan hasil dari *Software RapidMiner*



Gambar 4. Pohon Keputusan RapidMiner

3.4 Rule Hasil Dari Pengujian

1. Jika pelayanan respons karyawan baik maka konsumen merasa puas
2. Jika pelayanan respons karyawan cukup maka konsumen tidak puas
3. Jika pelayanan respons karyawan kurang maka konsumen tidak puas
4. Jika pelayanan respons karyawan sangat baik dan kemudahan menghubungi layanan sangat baik maka konsumen puas
5. Jika pelayanan respons karyawan sangat baik dan kemudahan menghubungi layanan baik maka konsumen tidak puas

Dari dua hasil pengujian yang telah dilakukan yaitu proses secara *manual* dan menggunakan *Software RapidMiner 5.3.000* dapat peneliti ambil sebuah kesimpulan bahwa hasil pengujian sangat baik karena *Rule* yang dihasilkan adalah sama. Algoritma C4.5 dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi dapat diperoleh dengan jelas, baik dalam struktur pohon keputusan maupun aturan *if-then*, sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan. Pada pemilihan atribut juga sangat mempengaruhi dalam pengolahan Algoritma C4.5 karena keputusan sangat bergantung pada atribut yang dipilih.

4. KESIMPULAN

Dari pembahasan di atas dapat ditarik kesimpulan sebagai berikut:

- a. Hasil dari pengolahan metode *Data Mining* dengan menggunakan algoritma C4.5 ternyata mampu mengklasifikasi tingkat kepuasan konsumen terhadap kualitas Pelayanan. Dengan nilai *presisi* prediksi puas sebesar 61.7% dan nilai *presisi* prediksi tidakpuas 38,3%.
- b. Pohon keputusan yang dihasilkan oleh teknik klasifikasi algoritma C4.5 dari variabel yang memiliki nilai *gain* tertinggi yaitu Pelayanan 2 yang merupakan respons karyawan saat menghubungi layanan konsumen (*responsiveness*). Sehingga perusahaan dapat menjadikan acuan untuk memperbaiki kualitas pelayanan untuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan.

REFERENCES

- [1] S. Takalapeta, "Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan," vol. 3, no. 3, pp. 34–38, 2018.
- [2] W. Insyroh and T. Setyowati, "Determinan Kualitas Pelayanan , Harga dan Produk Terhadap Kepuasan Pelanggan Telkomsel Flash Di Bondowoso," vol. 2, no. 3, pp. 132–140, 2018.
- [3] E. Elisa, "PREDIKSI PROFIT PADA PERUSAHAAN DENGAN," vol. 5, no. 2, pp. 179–189, 2018.
- [4] N. T. Nugroho, "PENGARUH KUALITAS PELAYANAN TERHADAP KEPUASAN DAN LOYALITAS PELANGGAN (Survey pada Pelanggan Speedy Telkom di Kota Surakarta) Novemy," vol. 12, no. 2, pp. 114–122, 2015.
- [5] I. R. Nasution, Dzukron Hamidan, "THE INFLUENCE OF SERVICE QUALITY TO CUSTOMER SATISFACTION FIRST MEDIA IN," vol. 4, no. 3, pp. 2332–2341, 2017.
- [6] S. Faisal, "Klasifikasi data minning menggunakan algoritma c4.5 terhadap kepuasan pelanggan sewa kamera cikarang," vol. 4, no. April, pp. 1–8, 2019.
- [7] I. P. Hariyadi, "DATA MINING ANALISA POLA PEMBELIAN PRODUK," pp. 6–7, 2016.
- [8] M. F. Rosas, "Classification and Regression Decision Tree: A Mining Technique for Students ' Insights on the University Services with Text Analysis," 2018 9th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf., no. 1, pp. 971–976, 2018.
- [9] A. S. Febriarini and E. Z. Astuti, "Penerapan Algoritma C4 . 5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang," pp. 95–103.