

## Penerapan Algoritma C4.5 Dalam Mengukur Kepuasan Pengunjung Terhadap Fasilitas di Taman Margasatwa Jakarta

Hendri<sup>1</sup>, Dony Oscar<sup>2</sup>

<sup>1</sup>STMIK Nusa Mandiri,

<sup>2</sup>Universitas Bina Sarana Informatika

e-mail: [hendri.hed@nusamandiri.ac.id](mailto:hendri.hed@nusamandiri.ac.id), [Dony.dor@bsi.ac.id](mailto:Dony.dor@bsi.ac.id)

Diterima	Direvisi	Disetujui
21-04-2021	01-06-2021	09-06-2021

**Abstrak** - Kepuasan pengunjung adalah suatu tingkatan dimana kebutuhan, keinginan dan harapan dari pengunjung dapat terpenuhi yang mengakibatkan pengunjung akan datang lagi (Zelmiati Zelmiati, Waryono Waryono, 2017). Taman margasatwa adalah salah satu objek wisata binatang yang paling banyak diminati masyarakat. Banyak hal yang harus diperhatikan dalam mengembangkan suatu objek wisata yaitu dengan memberikan sesuatu yang sesuai dengan yang dibutuhkan dan keinginan calon pengunjung. Berdasarkan fenomena tersebut peneliti menemukan bahwa jumlah pengunjung yang terus meningkat masih ada masalah yang dihadapi di Taman Margasatwa yaitu dari segi fasilitas wisata. Penelitian ini bertujuan untuk memberi rekomendasi kepuasan pengunjung taman margasatwa dengan menggunakan teknik datamining penerapan datamining C4.5. Sumber data diperoleh dari penyebaran kuesioner kepada pengunjung yang pernah berkunjung ke taman margasatwa kepada 35 orang. Atribut yang digunakan sebagai parameter penilaian kepuasan pengunjung antara lain: Paling Diminati, Fasilitas Semuanya, Tempat parkir, Keamanan, dan Kebersihan. Hasil pengolahan metode C4.5 dengan menggunakan Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki.

**Kata Kunci:** Kepuasan, Metode C4.5, Taman Margasatwa

**Abstrak** - *Visitor satisfaction is a level where the needs, wants and expectations of visitors can be met which results in visitors coming again (Zelmiati Zelmiati, Waryono Waryono, 2017). Wildlife parks are one of the most popular animal attractions in the community. There are many things that must be considered in developing a tourist attraction, namely by providing something that is in accordance with the needs and desires of potential visitors. Based on this phenomenon, the researchers found that the number of visitors who continue to increase there are still problems faced at the Wildlife Park, namely in terms of tourist facilities. Sources of data were obtained from distributing questionnaires to 35 people who had visited the wildlife park. The attributes used as parameters for assessing visitor satisfaction include: Most Interested, All Facilities, Parking, Security, and Cleanliness. The result of processing the C4.5 method using a decision tree is a predictive model using a tree structure or a hierarchical structure.*

**Keywords:** Satisfaction, Method C4.5, Wildlife Park

### PENDAHULUAN

Taman margasatwa adalah salah satu objek wisata binatang yang paling banyak diminati masyarakat. Banyak hal yang harus diperhatikan dalam mengembangkan suatu objek wisata yaitu dengan memberikan sesuatu yang sesuai dengan yang dibutuhkan dan keinginan calon pengunjung. Antusiasme pengunjung terhadap suatu objek wisata berdampak pada tingginya jumlah pengunjung. Tingginya jumlah pengunjung dan dalam hal tersebut membuat kebun binatang menjadi salah satu destinasi wisata yang sangat diminati di Jakarta sehingga mengharuskan pihak pengelola untuk memperhatikan kepuasan pengunjung. Kenaikan jumlah pengunjung mengidentifikasi adanya kepuasan berkunjung. Sebuah tempat yang

dapat memuaskan pengunjung memiliki citra yang baik dari pengunjung namun hal tersebut bertolak belakang dengan opini-opini mengenai Taman Margasatwa mengenai fasilitas di Taman Margasatwa yang dianggap masih kurang. Kondisi kandang yang kurang terawat dan tempat parkir yang membuat pengunjung merasa kurang puas (Marhanah & Wahadi, 2016)

Kepuasan pengunjung adalah suatu tingkatan dimana kebutuhan, keinginan dan harapan dari pengunjung dapat terpenuhi yang mengakibatkan pengunjung akan datang lagi (Zelmiati Zelmiati, Waryono Waryono, 2017). Banyak faktor yang mempengaruhi kepuasan pengunjung antara lain : faktor kebersihan, fasilitas, tempat parkir, keamanan yang ada pada tempat objek wisata.

Berdasarkan fenomena tersebut peneliti menemukan bahwa jumlah pengunjung yang terus meningkat masih ada masalah yang dihadapi di Taman Margasatwa yaitu dari segi fasilitas wisata, namun, pada saat ini pihak taman margasatwa sedang melakukan peningkatan terhadap fasilitas yang diberikan kepada pengunjung dengan tujuan dapat mendapatkan kepuasan pengunjung. Berkait dengan hasil yang didapatkan peneliti, peneliti tertarik untuk melakukan penelitian lebih lanjut mengenai pengaruh fasilitas terhadap kepuasan pengunjung di Taman Margasatwa. Adapun judul dari penelitian yang oleh peneliti adalah "Penerapan Data Mining Menggunakan Metode Algoritma C4.5 Untuk Tingkat Kepuasan Pengunjung Terhadap Fasilitas Di Taman Margasatwa (Marhanah & Wahadi, 2016).

## KAJIAN LITERATUR

### A. Data Mining

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis yang menemukan keteraturan ini yaitu proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data besar (Junia et al., 2021). Menurut Yuli Mardi Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Mardi, 2017). Menurut para ahli data mining adalah proses menemukan pola pengetahuan dari data dalam jumlah besar. Sumber data dapat berupa database, data warehouses, Website, repositori, atau data yang disimpan ke sistem secara dinamis (Han et al., 2014)

### B. Pohon Keputusan

Pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan (Novianti et al., 2016).

Pohon keputusan adalah kemampuannya untuk mem-break down proses pengambilan keputusan akan lebih menginterpretasikan solusi dari permasalahan (Syarif Mayron Turnip, 2018). Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel input dengan sebuah variabel target (Firmansyah & AUFANY, 2016).

### C. Algoritma C4.5

Algoritma C4.5 merupakan kelompok algoritma pohon keputusan (Decision Tree).

Algoritma ini mempunyai input berupa training samples dan samples. Training samples berupa data yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya (Junia et al., 2021).

## METODE PENELITIAN

Metodologi adalah suatu cara yang digunakan dalam mendapatkan berbagai data yang akan diproses menjadi sebuah informasi yang lebih akurat sesuai dengan permasalahan yang akan diteliti. Metodologi penelitian juga dilakukan dengan mendeskripsikan masalah untuk memudahkan dalam memahami tahapan penelitian

### A. Pengumpulan Data

Berikut adalah cara yang dilakukan dalam pengumpulan data sebagai berikut :

#### 1. Pengumpulan Data Primer

Dilakukan untuk mendapatkan informasi yang digunakan peneliti dalam pembuatan penelitian agar hasil yang didapatkan akurat. Penelitian dilakukan dengan cara kuesioner dengan menggunakan google formulir

#### 2. Pengumpulan data sekunder

Pengumpulan data sekunder diperoleh melalui jurnal, makalah, laporan dll

### B. Populasi

Populasi adalah wilayah generalisasi yang terdiri atas obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik (Vogt, 2015). Populasi dalam penelitian ini adalah para responden yang pernah berkunjung ke Taman Margasatwa.

Dalam penelitian ini pengambilan sampel adalah menyebar kuesioner kepada masyarakat yang pernah berkunjung ke Taman Margasatwa. Pada penelitian ini jumlah populasi tidak diketahui sehingga sampel yang diambil dari kuesioner yang diisi oleh responden sebanyak 35 responden agar mewakili populasi tidak secara keseluruhan.

### C. Metode Analisis Data

Untuk menentukan Kepuasan Pengunjung Terhadap Fasilitas Taman Margasatwa Menggunakan Metode Algoritma C4.5 maka data yang akan digunakan dan di analisa data kuesioner responden yang pernah ke taman margasatwa dimana diperoleh dari pengisian kuesioner dari pengunjung. Untuk menentukan kepuasan pengunjung yaitu dibedakan menjadi dua kelas yaitu Puas dan Tidak Puas.

**D. Menentukan Atribut**

Tahap utama pada penelitian ini adalah menentukan jenis atribut yang nantinya akan dijadikan proses dalam menentukan kepuasan pengunjung dalam penelitian jumlah kriteria yang disediakan sebanyak 5 kriteria yaitu Paling Diminati, Fasilitas Semua, Tempat Parkir, Keamanan, Kebersihan

**E. Algoritma C4.5**

Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari impurity dan homogeneity dari kumpulan(Arifin & Fitriana, 2018), . Dari nilai Entropy tersebut kemudian dihitung nilai information gain masing-masing atribut. Penghitungan nilai Entropy digunakan rumus seperti dalam Persamaan

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

Berikut keterangannya :

S : Himpunan kasus

N : Jumlah partisi

Pi : Jumlah kasus pada partisi ke-*i*”

Information Gain adalah informasi yang didapatkan dari perubahan entropy pada suatu kumpulan data, baik melalui observasi atau bisa juga disimpulkan dengan cara melakukan partisipasi terhadap suatu set data(Arifin & Fitriana, 2018).

$$GAIN(S, A) = Entropy(S) - \sum_{i=1}^n |S_i| \cdot \log_2(|S_i|) \dots \dots \dots (2)$$

*Entropy(S<sub>i</sub>)* .....(2)

Berikut keterangannya :

S : Himpunan kasus

N : Jumlah partisi atribut A

|S<sub>i</sub>| : Jumlah kasus pada partisi ke-*i*

|S| : Jumlah kasus da

**HASIL DAN PEMBAHASAN**

Data yang peneliti gunakan dalam penelitian ini adalah data kuesioner. Data diperoleh dari pembagian kuesioner sebanyak 35 responden. Data yang dikumpulkan yaitu data kuesioner yang dibagikan kepada masyarakat yang pernah berkunjung ke Taman Margasatwa. Penyebaran kuesioner dimulai pada bulan April 2020. Untuk mempermudah pengisian dan pengolahan data ,maka dibuatkanlah kriteria dan alternatif pada kepuasan pengunjung kriteria yang digunakan dalam penelitian ini yaitu wisata yang paling diminati,fasilitas yang ada,tempat parkir,keamanan,kebersihan

**1. Data Training**

Data training adalah data yang digunakan untuk melakukan perhitungan menggunakan metode

Algoritma C4.5 berikut adalah tabel dari data training yang digunakan

Tabel 1. Data Training

N o.	Paling Dinik mati	Fasilitas Semua	Tempat Parkir	Keamanan	Kebersihan	Hasil
1	4	3	3	2	2	Tidak Puas
2	4	4	2	4	4	Puas
3	3	4	2	3	3	Puas
4	4	3	2	3	2	puas
5	3	3	3	3	3	puas
6	3	4	2	3	3	puas
7	3	1	2	2	1	Tidak Puas
8	3	3	3	3	3	Puas
9	3	3	3	3	4	Puas
10	3	4	2	3	3	Puas
11	3	3	3	3	3	Puas
12	3	4	3	4	3	puas
13	3	3	3	3	3	Puas
14	3	3	3	4	4	Puas
15	3	3	2	3	3	Tidak Puas
16	3	4	3	4	4	Puas
17	4	4	3	3	3	Puas
18	3	2	2	1	2	Tidak Puas
19	4	4	3	3	3	Puas
20	3	3	2	3	3	Tidak Puas
21	3	4	2	4	4	Tidak Puas
22	3	3	3	3	3	Puas
23	3	3	2	1	2	Tidak Puas
24	3	3	2	3	4	Tidak Puas
25	3	3	3	3	3	Puas
26	4	4	3	3	4	Puas
27	4	4	3	4	4	Puas
28	3	3	3	4	3	Puas
29	3	3	2	3	2	Tidak Puas
30	4	3	3	3	4	Puas
31	3	3	2	2	3	Tidak Puas
32	3	3	3	3	3	Puas
33	3	3	3	2	3	Tidak Puas
34	3	4	3	4	3	Puas
35	3	4	3	2	3	Tidak Puas

2. Data Testing

Berdasarkan pada data training diatas,data yang diuji dengan menggunakan Data testing Algoritma C4.5 dalah seperti tabel dibawah ini

Tabel 2. Data Testing

No .	Paling Dinikmati	Fasilitas Semua	Tempat Parkir	Keamanan	Kebersihan	Hasil
1	4	3	3	2	2	Tidak Puas
2	4	4	2	4	4	Puas
3	3	4	2	3	3	Puas
4	4	3	2	3	2	puas
5	3	3	3	3	3	puas
6	3	4	2	3	3	puas
7	3	1	2	2	1	Tidak Puas
8	3	3	3	3	3	Puas
9	3	3	3	3	4	Puas
10	3	4	2	3	3	Puas
11	3	3	3	3	3	Puas
12	3	4	3	4	3	puas
13	3	3	3	3	3	Puas
14	3	3	3	4	4	Puas
15	3	3	2	3	3	Tidak Puas
16	3	4	3	4	4	Puas
17	4	4	3	3	3	Puas
18	3	2	2	1	2	Tidak Puas
19	4	4	3	3	3	Puas
20	3	3	2	3	3	Tidak Puas
21	3	4	2	4	4	Tidak Puas
22	3	3	3	3	3	Puas
23	3	3	2	1	2	Tidak Puas
24	3	3	2	3	4	Tidak Puas
25	3	3	3	3	3	Puas
26	4	4	3	3	4	Puas
27	4	4	3	4	4	Puas
28	3	3	3	4	3	Puas
29	3	3	2	3	2	Tidak Puas
30	4	3	3	3	4	Puas

a. Perhitungan Entropy & Gain

Berikut ini adalah cara perhitungan entropy total dan gain dari paling diminati yaitu sebagai berikut :

Tabel 3. Perhitungan Node 1

No de		Jumlah ( S )	Puas ( Si )	Tidak ( Si )	Entropy	Gain
1	Total	30	21	9	0,881290899	
	Paling Diminati					0,042856158
	Sangat Setuju	8	7	1	0,543564443	
	Setuju	22	14	8	0,945660305	
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
	Fasilitas Semua					0,189363821
	Sangat Setuju	11	10	1	0,439496987	
	Setuju	17	11	6	0,936667382	
	Tidak Setuju	1	0	1	0	
	Sangat Tidak Setuju	1	0	1	0	
	Tempat Parkir					0,281859427
	Sangat Setuju	0	0	0	0	
	Setuju	17	16	1	0,322756959	
	Tidak Setuju	13	5	8	0,961236605	
	Sangat Tidak Setuju	0	0	0	0	
	Keamanan					0,27299179
	Sangat Setuju	7	6	1	0,591672779	
	Setuju	19	15	4	0,74248757	
	Tidak Setuju	2	0	2	0	
	Sangat Tidak Setuju	2	0	2	0	
	Kebersihan					0,248453445
	Sangat Setuju	9	7	2	0,764204507	
	Setuju	15	13	2	0,566509507	
	Tidak Setuju	5	1	4	0,721928095	
	Sangat Tidak Setuju	1	0	1	0	

Baris total kolom Entropy pada Tabel 3 dihitung dengan rumus sebagai berikut :

$$Entropy ( Total ) = \left(-\frac{21}{30} * \log_2\left(\frac{21}{30}\right)\right) + \left(-\frac{9}{30} * \log_2\left(\frac{9}{30}\right)\right)$$

$$Entropy ( Total ) = 0,881290899$$

Sementara itu nilai Gain pada baris Paling Diminati dihitung dengan menggunakan rumus Gain Sebagai berikut :

$$Gain ( Total, Paling Diminati ) = Entropy ( Total )$$

$$\sum_{i=1}^n \frac{cuaca}{l_{Total} i} \times Entropy ( Paling Diminati )$$

$$\text{Gain ( Total Paling Diminati )} = 0,881290899 - \left( \frac{17}{30} * \right) + \left( \frac{22}{30} * 0,94566030 \right) + \left( \frac{0}{30} * 0 \right) + \left( \frac{0}{30} * 0 \right)$$

$$\text{Gain ( Total Paling Diminati )} = 0.042$$

b. Implementasi Klasifikasi Algoritma C4.5 Menggunakan Rapid Miner

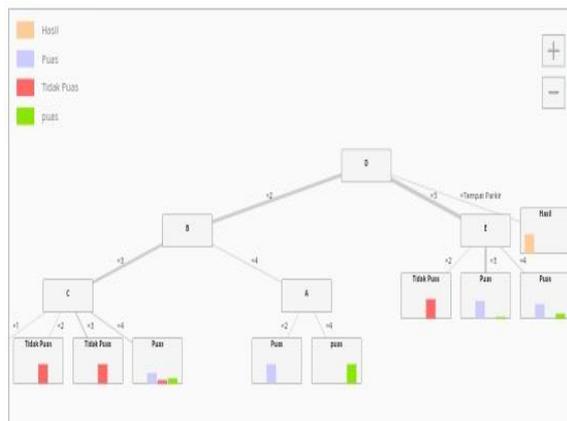
Berdasarkan padari Tabel diatas data yang diuji adalah 30 data kepuasan Pengunjung dan akan diuji menggunakan Tools rapid miner Sehingga didapatkan hasil klasifikasi kepuasan pengunjung. Hasil dari proses pertama maka akan mendapatkan hasil dalam ExampleSet yang dapat dilihat pada gambar 3 Hasilnya berupa informasi mengenai perbandingan dari kepuasan pengunjung pada data terhadap kepuasan oleh Tolss RapidMiner.

	I	J	K	L	M	N	O
	Categorical	Categorical	Numerical	Categorical	Categorical	Categorical	Categorical
6	Paling Diminati	?	?	?	?	?	0.04285615767...
7	?	Sangat Setuju	8.0	7	1	0.54356444319...	?
8	?	Setuju	22.0	14	8	0.04566030450...	?
9	?	Tidak Setuju	0.0	0	0	0	?
10	?	Sangat Tidak Se...	0.0	0	0	0	?
11	Fasilitas Semsuah	?	?	?	?	?	0.18936382096...
12	?	Sangat Setuju	11.0	10	1	0.491019698692...	?
13	?	Setuju	17.0	11	6	0.9366738187...	?
14	?	Tidak Setuju	1.0	0	1	0	?
15	?	Sangat Tidak Se...	1.0	0	1	0	?
16	Tempat Parkir	?	?	?	?	?	0.2818942714...

Gambar 1. Hasil ExampleSet Pada Tools RapidMiner

c. Hasil Pohon Keputusan Yang Dihasilkan

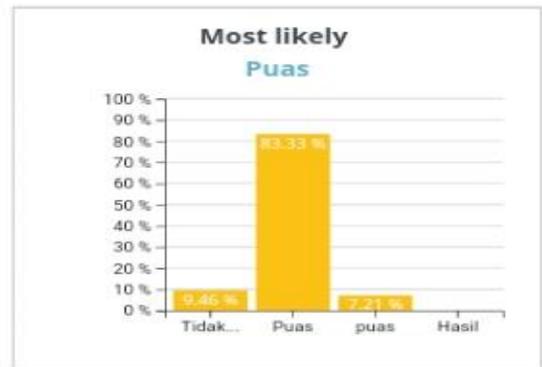
Setelah dilakukan perhitungan Algoritma c4.5 pada RapidMiner, Maka terbentuk pohon keputusan seperti gambar dibawah ini:



Gambar 2. Pohon Keputusan

4.4 Hasil Penelitian

Setelah melakukan perhitungan dari hasil kuesioner dan kemudian dihitung, maka dari keseluruhan perhitungan tersebut diperoleh hasil sebagai berikut:



Gambar 3. Hasil

Hasil Perhitungan

Potensi dari penerapan data mining menggunakan metode Algoritma C4.5 dan tools Rapid Miner pada pengelolaan data kepuasan pengunjung taman margasatwa menghasilkan potensi yang baik. Dibuktikan dengan hasil 83% puas terhadap fasilitas yang ada ditaman margasatwa

KESIMPULAN

Dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan bahwa kepuasan pengunjung taman margasatwa dapat diprediksi dan dievaluasi dengan memanfaatkan Algoritma C4.5 untuk memprediksi kepuasan pengunjung dengan traning yang diperoleh. Dari Penelitian diatas dapat disimpulkan bahwa 83% pengunjung puas terhadap fasilitas ditaman margasatwa.

REFERENSI

Arifin, M. F., & Fitriana, D. (2018). Penerapan Algoritma Klasifikasi C4.5 Dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus: PT Atria Artha Persada. *InComTech*, 8(2), 87–102. <https://doi.org/10.22441/incomtech.v8i1.2198>

Firmansyah, M., & Aufany, R. (2016). Implementasi Metode Decision Tree Dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank. *Infokam*, XII(1), 1–12.

Han, J., Kamber, M., & Pei, J. (2014). Data mining: Data mining concepts and techniques. In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. morgan kaufmann. <https://doi.org/10.1109/ICMIRA.2013.45>

Junia, A., Studi, P., Informatika, T., Pelita, S., Ji, N., Iskandar, S., & No, M. (2021). *Implementasi Algoritma C4 . 5 Untuk Mengukur Tingkat Kepuasan BPJS Katernaga Kerjaan*. 2(2), 12–15.

- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- Marhanah, S., & Wahadi, W. H. (2016). Pengaruh Fasilitas Wisata Dan Kualitas Pelayanan Terhadap Kepuasan Pengunjung Di Taman Margasatwa Ragunan Jakarta. *Jurnal Manajemen Resort Dan Leisure*, 13(1). <https://doi.org/10.17509/jurel.v13i1.2134>
- Novianti, B., Rismawan, T., & Bahri, S. (2016). Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak). *Jurnal Coding, Sistem Komputer Untan*, 04(3), 75–84.
- Syarif Mayron Turnip, P. S. (2018). Analisis Pola Penyebaran Penyakit dengan. 03(479), 3–7. <https://media.neliti.com/media/publications/282579-analisis-pola-penyebaran-penyakit-dengan-767de6bd.pdf>
- Vogt, W. (2015). *Proportional Stratified Random Sample*. Dictionary of Statistics & Methodology. <https://doi.org/10.4135/9781412983907.n1534>
- Zelmiati Zelmiati, Waryono Waryono, Y. A. (2017). PENGARUH KUALITAS PELAYANAN TERHADAP KEPUASAN PENGUNJUNG DI OBJEK WISATA MINANG FANTASI KOTA PADANG PANJANG. *E-Journal Home Economic and Tourism*, 14. <http://ejournal.unp.ac.id/index.php/jhet/article/view/7180>