

ANALISA KEPUASAN PENGUNJUNG MENGGUNAKAN METODE *DECISION TREE* DI PERKAMPUNGAN BUDAYA BETAWI SETU BABAKAN JAKARTA

Kudiantoro Widiyanto¹, Achmad Rifai², Mutiara Nur Pratiwi³

¹Sistem Informasi Akuntansi

^{2,3}Sistem Informasi

¹ Universitas Bina Sarana Informatika, Jl. Kamal Raya No. 18, Ringroad Barat Cengkareng Jakarta

^{2,3} STMIK Nusa Mandiri, Jl. Kramat Raya No. 18 Jakarta

Email: kudiantoro.kdw@bsi.ac.id, achmad.acf@nusamandiri.ac.id, mutiaranp1105@nusamandiri.ac.id

Abstract

Tourist attraction managers in Jakarta make various efforts to win the competition, among others by increasing tourist visitor satisfaction. One of the attractions in Jakarta is Setu Babakan which is one of the attractive tourist attractions for tourists who want to enjoy the typical atmosphere of the countryside or witness native Betawi culture firsthand. In Setu Babakan, visitors can see natural scenery, enjoy typical Betawi culinary which is relatively cheap, and watch Betawi cultural arts performances. In addition, the manager of Setu Babakan also provides various facilities such as toilets, places of worship, children's playgrounds, and others to support visitors' activities. To visit Setu Babakan, there are many public vehicles that pass through this village and the cost is quite cheap. This study aims to analyze the visitors' satisfaction of Setu Babakan by using the attributes of cost, facilities, accessibility, and service. Researchers use the decision tree method C4.5 algorithm. From the decision tree image processed by RapidMiner it can be seen that the factors of visitor satisfaction are facilities and accessibility. After that the trial was performed on the dataset 3 times from the questionnaire data on 50 visitors obtained an accuracy of 100%. The classification results with the C4.5 algorithm of 100% are considered appropriate to be used to measure the level of visitor satisfaction with costs, facilities, accessibility, and service.

Keywords: *Decision Tree, C4.5 Algorithm, Visitor Satisfaction*

Abstrak

Pengelola objek wisata di Jakarta melakukan berbagai upaya untuk memenangkan persaingan, antara lain dengan meningkatkan kepuasan pengunjung wisata. Salah satu objek wisata yang berada di Jakarta adalah Setu Babakan yang merupakan salah satu objek wisata menarik bagi wisatawan yang ingin menikmati suasana khas pedesaan atau menyaksikan budaya Betawi asli secara langsung. Di Setu Babakan pengunjung dapat melihat pemandangan alam, menikmati kuliner khas betawi yang harganya relatif murah, dan menyaksikan pagelaran seni budaya betawi. Selain itu, pihak pengelola Setu Babakan juga menyediakan berbagai fasilitas seperti toilet, tempat ibadah, tempat bermain anak, dan lain-lain guna menunjang aktivitas pengunjung. Untuk berkunjung ke Setu Babakan, terdapat banyak kendaraan umum yang melewati perkampungan ini dan biaya yang cukup murah. Penelitian ini bertujuan untuk menganalisis kepuasan pengunjung Setu Babakan dengan menggunakan atribut biaya, fasilitas, aksesibilitas, dan pelayanan. Peneliti menggunakan metode *decision tree* algoritma C4.5. Dari gambaran *decision tree* yang diolah dengan RapidMiner dapat diketahui bahwa faktor kepuasan pengunjung adalah fasilitas dan aksesibilitas. Setelah itu dilakukan ujicoba pada dataset sebanyak 3 kali dari data kuesioner terhadap 50 orang pengunjung diperoleh akurasi sebesar 100%. Hasil klasifikasi dengan algoritma C4.5 sebesar 100% dianggap tepat digunakan untuk mengukur tingkat kepuasan pengunjung terhadap biaya, fasilitas, aksesibilitas, dan pelayanan.

Kata kunci : *Decision Tree, Algoritma C4.5, Kepuasan Pengunjung*

1. Pendahuluan

Provinsi DKI Jakarta merupakan ibukota negara Indonesia yang dijadikan pusat pemerintahan dan perekonomian. Selain itu Jakarta juga memiliki potensi sebagai destinasi wisata. Potensi wisata yang ada di Jakarta antara lain wisata budaya, wisata religi, wisata minat khusus, minat sejarah, wisata kuliner dan wisata belanja. Berdasarkan data Dinas Pariwisata dan Kebudayaan Provinsi DKI Jakarta tahun 2017,

obyek wisata di DKI Jakarta berjumlah 17 obyek wisata, dan sebagian besar obyek wisata di Jakarta merupakan obyek wisata budaya. Budaya yang menjadi ciri khas suatu daerah menjadi daya tarik tersendiri untuk dikembangkan menjadi daerah tujuan wisata sehingga dapat menarik wisatawan lokal maupun mancanegara.

Setu Babakan merupakan sebuah kawasan perkampungan yang ditetapkan Pemerintah Jakarta sebagai tempat pelestarian dan pengembangan budaya Betawi secara

berkesinambungan. Perkampungan yang terletak di Srengseng Sawah, kecamatan Jagakarsa, Jakarta Selatan ini merupakan salah satu objek wisata yang menarik bagi wisatawan yang ingin menikmati suasana khas pedesaan atau menyaksikan budaya Betawi asli secara langsung. Agar dapat bersaing dengan objek wisata lain, sebuah kawasan wisata harus dapat memuaskan pengunjung lama dan menciptakan pengunjung baru [1].

Dalam hal ini pihak pengelola Setu Babakan harus bisa mengetahui variabel apa saja yang menghasilkan hubungan sebab akibat dalam mengklasifikasikan pengunjung puas atau tidak puas. Untuk mengatasi hal tersebut diperlukan sebuah metode yang dapat digunakan untuk mengetahui variabel apa saja yang mempengaruhi kepuasan pengunjung Setu Babakan. Dalam hal ini, penulis menggunakan algoritma data mining dengan metode *decision tree* menggunakan algoritma C4.5. Menurut Berry & Linoff Algoritma C4.5 merupakan salah satu solusi pemecahan kasus yang sering digunakan dalam pemecahan masalah pada teknik klasifikasi [2]. Keluaran dari algoritma C4.5 itu berupa sebuah *decision tree* layaknya teknik klasifikasi lain.

Menurut Han & Kamber, *decision tree* adalah sebuah diagram alir yang mirip dengan struktur pohon, dimana setiap internal node menotasikan atribut yang diuji, setiap cabangnya merepresentasikan hasil dari atribut tersebut dan leaf node merepresentasikan kelas-kelas tertentu [3].

Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagikan kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain.

Kepuasan pelanggan adalah tingkat perasaan seseorang setelah membandingkan kinerja atau hasil yang ia rasakan dibandingkan dengan harapannya. Kepuasan pelanggan berjalan berbarengan dengan ketidakpuasan [4]. Artinya respon pelanggan terhadap evaluasi dan ketidaksesuaian yang dirasakan antara harapan sebelumnya dan kinerja aktual produk yang dirasakan setelah pemakaian.

2. Metode

a. Metode Pengumpulan Data

Penelitian dilakukan pada bulan April sampai dengan bulan Mei 2019 di Setu Babakan. Lokasi ini dipilih karena Setu Babakan merupakan salah satu objek wisata budaya yang ada di Jakarta. Penelitian ini bertujuan untuk mengetahui tingkat kepuasan pengunjung yang datang ke Setu Babakan. Pada penelitian ini, peneliti membagikan 50 kuesioner kepada pengunjung Setu Babakan. Atribut kuesioner terdiri dari nama, alamat, biaya, fasilitas, aksesibilitas, pelayanan, serta hasil puas atau tidak puasnya pengunjung Setu Babakan.

b. Metode Analisis Data

Data yang digunakan dalam penelitian ini merupakan data yang diperoleh berdasarkan survei yang dilakukan di Setu Babakan dengan membagikan 50 kuesioner kepada pengunjung. Dengan urutan analisa data meliputi : 1) Peneliti membagikan kuesioner kepada 50 responden yaitu pengunjung Setu Babakan; 2) Mengumpulkan kembali kuesioner yang telah diisi oleh responden, kemudian melakukan penghitungan untuk mengetahui jumlah puas dan tidak puas pada data kuesioner pengunjung Setu Babakan; 3) Data selection, memilih data yang akan digunakan pada penelitian; 4) Data cleaning, membersihkan atribut yang hilang atau kosong dan merubah data yang tidak konsisten; 5) Menghitung data dengan rumus *decision tree* pada algoritma C4.5.; 6) Mengolah data dengan Rapid Miner; 7) Melakukan pengujian terhadap hasil penghitungan Rapid Miner dengan pembagian data; 8) *training* dan data *testing* yang berbeda; 9) Hasil pengolahan data menghasilkan keputusan pengunjung Setu Babakan; 10) Melakukan validasi dan evaluasi data menggunakan teknik yang ada dalam Rapid Miner; 11) Mengimplementasikan rule dengan membuat program dari data yang telah diolah.

3. Hasil Dan Pembahasan

a. Pengolahan Data

Data yang digunakan dalam penelitian ini diperoleh dari data kuesioner sebanyak 50 kuesioner. Peneliti membagikan kuesioner kepada pengunjung Setu Babakan pada bulan April- Mei 2019 dengan atribut nama, alamat, biaya, fasilitas, aksesibilitas, pelayanan, dan kepuasan.

Data kuesioner diolah dengan Microsoft Office Excel sebagai berikut ini :

A1	No															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	No	Nama	Alamat	x1-Biaya	Kata rata	x2-Fasilitas	Kata rata	x3-Aksesibilitas	Kata rata	x4-Pelayanan	Kata rata	Hasil				
2				x1	x1	x2	x2	x3	x3	x4	x4					
3	1	Nina	Cipodale	1	1	3	2	2,2	3	1	1	3	Puas			
4	2	Dewi Haryati	Pasar Minasa	5	5	4	3	3,5	3	3	3	3	Puas			
5	3	Muhammad Sai	Pasar Kobo	1	1	3	3	3	1	1	1	3	Puas			
6	4	Aisyah	Lantana Amara	4	4	4	2	3	3	3	4	4	Puas			
7	5	Idayanti	Dipole	2	2	3	2	2,2	2	3	2	2	Tidak Puas			
8	6	Abdul Aziz	Jasikara	3	3	2	2	2	3	3	4	3	Tidak Puas			
9	7	Masacroh	Pasar Kobo	1	1	1	3	3,2	1	1	1	1	Puas			
10	8	Aliya Putri	Pondole Labu	5	5	3	3	3	3	4	4	4	Puas			
11	9	Maulana	Legal Parang	3	3	2	2	2	3	3	2	2	Tidak Puas			
12	10	Ariyanti	Tawal Pacane	4	4	4	2	3	4	4	3	4	Puas			
13	11	Muh. Akbar	Cilandak	3	3	2	2	2	1	3	2	2	Tidak Puas			
14	12	Sigit Santosa	Cipodale	4	4	4	3	3,5	4	4	3	4	Puas			
15	13	Faridah	Mempang pratap	1	1	3	3	3	1	1	3	3	Puas			
16	14	Taliesari	Taliesari	5	5	4	3	3,5	4	4	4	4	Puas			
17	15	Abdul Salur	Dipole	3	3	3	3	3	3	3	2	3	Puas			
18	16	Heriyanto	Cinca	3	3	3	2	2,5	4	3	2	3	Tidak Puas			
19	17	Siska Andriyani	Pasar Minasa	1	1	1	3	3,2	2	1	1	3	Puas			
20	18	Maulidia Nur	Lantana Amara	4	4	3	2	2,5	4	4	4	4	Puas			
21	19	Galih Suputra	Jalagadung	3	3	1	3	3,2	2	3	1	3	Puas			
22	20	Novianti	Talipadara	3	3	3	2	2,5	3	3	3	3	Tidak Puas			
23	21	Riska Febriani	Kubugutan	1	1	1	3	3,2	1	1	1	3	Puas			
24	Sheet1															

Sumber Penelitian 2019

Gambar 1. Data Excel Kuesioner

Setelah data dimasukan atau diolah menggunakan Microsoft Office Excel, maka dilakukan pengolahan data kuesioner,dengan cara mengambil data rata-ratax1(Biaya), rata-ratax2

(Fasilitas), rata-rata x3 (Aksesibilitas), rata-rata x4 (Pelayanan) dan Y (Hasil) sehingga hasilnya akan seperti pada tabel berikut ini:

No	Nama	Biaya (X1)	Fasilitas (X2)	Aksesibilitas (X3)	Pelayanan (X4)	Hasil (Y)
1	Nina	4	2,50	3,50	4	Puas
2	Dewi Haryati	5	3,50	3,00	3	Puas
3	Muhammad Sai	4	3,00	3,75	4	Puas
4	Aisyah	4	3,00	3,50	3	Puas
5	Ikayanti	2	2,50	2,25	3	Tidak Puas
6	Abdul Aziz	3	2,00	3,25	3	Tidak Puas
7	Masacroh	1	3,50	1,00	1	Puas
8	Aliya Putri	5	3,00	3,75	3	Puas
9	Maulana	3	2,00	2,50	3	Tidak Puas
10	Ariyanti	4	3,00	3,75	4	Puas
11	Muh. Akbar	3	2,00	2,75	3	Tidak Puas
12	Sigit Santosa	4	3,50	3,75	4	Puas
13	Faridah	4	3,00	3,50	3	Puas
14	Taliesari	5	3,50	1,00	1	Puas
15	Abdul Salur	3	3,00	2,75	3	Puas
16	Heriyanto	3	2,50	3,00	3	Tidak Puas
17	Siska Andriyani	4	3,50	3,25	3	Puas
18	Maulidia Nur	4	2,50	4,00	3	Puas
19	Galih Saputra	3	3,50	3,25	3	Puas
20	Novianti	3	2,50	3,00	3	Tidak Puas
21	Riska Febriani	4	3,50	3,75	4	Puas

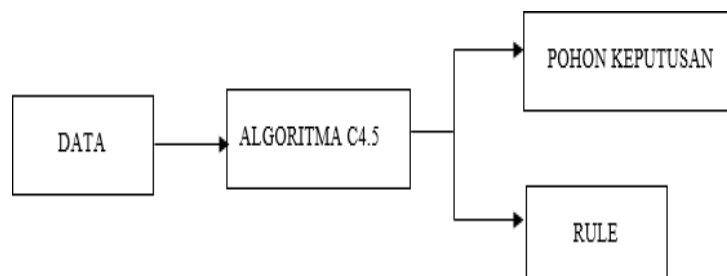
Sumber Penelitian 2019

Gambar 2. Data kuesioner yang telah diolah

Dari 50 kuesioner diperoleh hasil 36 kuesioner yang menyatakan PUAS dan 14 kuesioner menyatakan TIDAK PUAS.

b. Proses Data Mining

Dalam tahap ini data akan diproses menggunakan Rapid Miner dan pengujiannya menggunakan *cross validation*. Maka data akan terlihat nilai akurasi, pohon keputusan, danrulenya.



Gambar 3. Alur pengolahan data

Berikut adalah data yang akan di import ke dalam RapidMiner :

Import Data - Format your columns.

Format your columns.

Date format: MMM d, yyyy h:mm:ss z

☐ Replace errors with missing values

	Nama	Biaya (X1)	Fasilitas (X2)	Aksesibilita...	Pelayanan (...)	Hasil (Y)
	polynomial	integer	real	real	integer	polynomial (label)
1	Nina	4	2.500	3.500	4	Puas
2	Dewi Hartiyali	5	3.500	3.000	3	Puas
3	Muhammad Sari	4	3.000	3.750	4	Puas
4	Aisyah	4	3.000	3.500	3	Puas
5	Idiyanti	2	2.500	2.250	3	Tidak Puas
6	Abdul Aziz	3	2.000	3.250	3	Tidak Puas
7	Masruri	4	2.500	4.000	4	Puas
8	Alya Putri	5	3.000	3.750	3	Puas
9	Maulana	3	2.000	2.500	3	Tidak Puas
10	Anisanti	4	3.000	3.750	4	Puas
11	Muh. Akbar	3	2.000	2.750	3	Tidak Puas
12	Siti Santosa	4	2.500	3.750	4	Puas

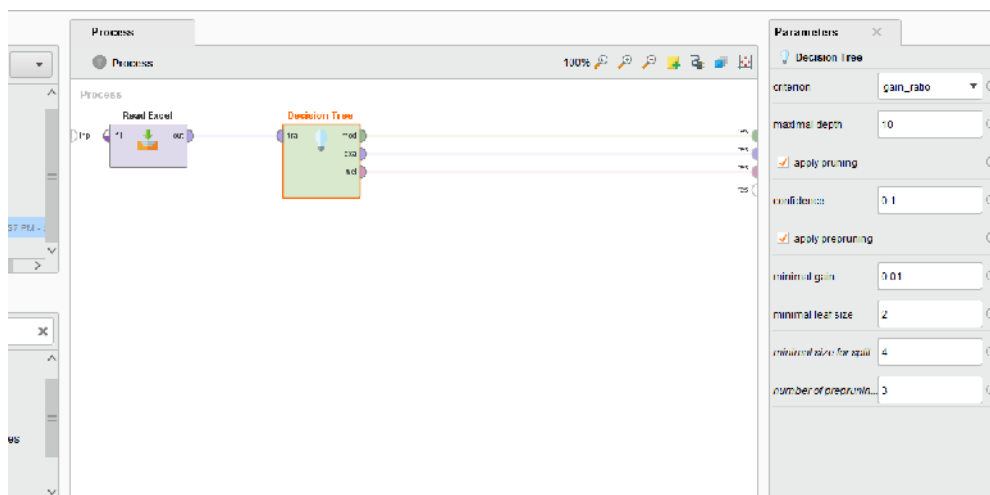
no problems.

Previous Finish Cancel

Sumber Penelitian 2019

Gambar 4.
Import Data ke RapidMiner

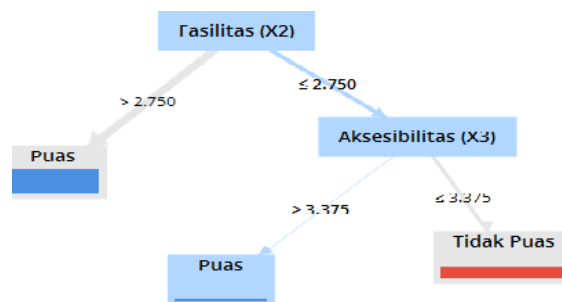
Berikutnya data diolah dengan menggunakan keputusan dan rule.
RapidMiner untuk dapat mengetahui pohon



Sumber Penelitian 2019

Gambar 5.
Proses pengolahan data

Setelah penghitungan dengan algoritma C4.5 menggunakan RapidMiner, maka akan terbentuk sebuah pohon keputusan seperti dibawah ini :



Sumber Penelitian 2019

Gambar 6.
Decision Tree

Dari gambaran *decision tree* dapat disimpulkan bahwa faktor kepuasan pengunjung adalah

fasilitas dan aksesibilitas. *Decision tree* menghasilkan rule sebagai berikut :

Tabel 1.
Tabel Rule

Rule untuk hasil = PUAS
1. IF Fasilitas >2.75 THEN PUAS
2. IF Fasilitas <=2.75 AND Aksesibilitas >3.375 THEN PUAS
Rule untuk hasil = TIDAK PUAS
1. IF Fasilitas <=2.75 AND Aksesibilitas <=3.375 THEN TIDAK PUAS

Pengujian Rules Terhadap Data Kepuasan Pengunjung dengan Variabel Data Kuesioner

Pengujian ini dilakukan untuk mengetahui kinerja algoritma C4.5 dalam melakukan klasifikasi terhadap kelas yang telah ditentukan dalam uji coba ini. Pengujian ini dilakukan sebanyak 3 kali dengan menggunakan *K-Fold Validation* yang berbeda pada operator *cross validation*. *K-Fold validation* berfungsi untuk membagi jumlah data training dan data testing pada data yang diuji.

Pengujian terdiri dari: *K-Fold validation* 10; *K-Fold validation* 6; *K-Fold validation* 4.

Algoritma C4.5 melakukan *training* terhadap data-data yang telah dibagi oleh *cross validation* menjadi dua kategori yaitu *training* dan *testing*. *Training* terdiri dari *decision tree*, sedangkan *testing* terdiri dari *apply model* dan *performance*.

Pengujian K-Fold Validation 10

Pengujian ini dilakukan dengan membagi sebanyak 10 bagian pada 50 data kuesioner yang akan diuji. 10 bagian tersebut terdiri dari 9 bagian data

training dan 1 bagian data *testing*. Dan dalam pengujian ini akan menghasilkan nilai akurasi, *precision*, *recall* dan *AUC* (Area Under Curve). Hasil akurasi *confusion matrix* pada pengujian 1 :

Tabel 2.

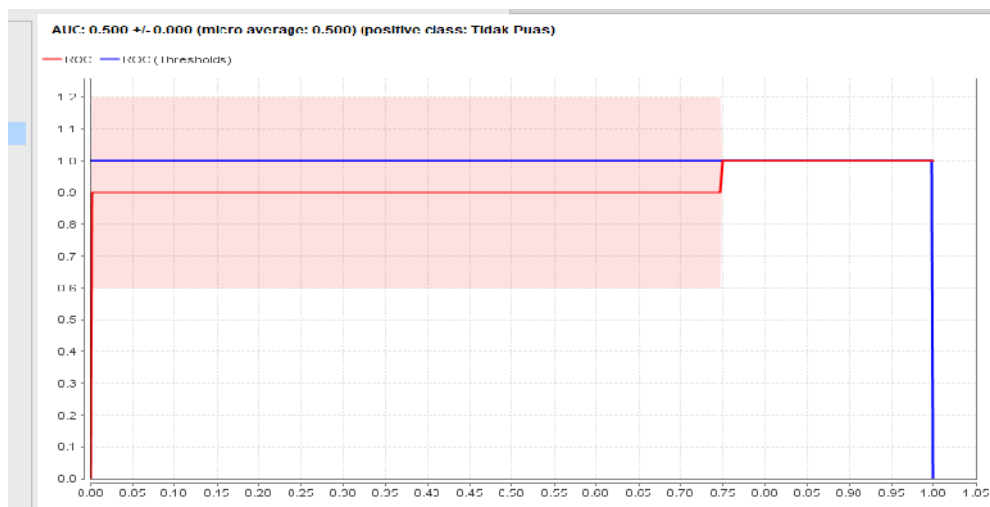
Tabel *confusion matrix* pengujian 1

	Actual	
	Puas	Tidak Puas
Pred. Puas	32	0
Pred. Tidak Puas	3	15

Jumlah true positive (tp) 32 data, false positive (fp) 3 data, true negative (tn) 15 data, dan false negative 0 data. Dari hasil pengujian 1 pada Rapid Miner diperoleh akurasi sebesar 94%. Berikut adalah perhitungan akurasi:

$$\begin{aligned}
 \text{Akurasi} &= \frac{tp+tn}{tp+tn+fp+fn} \times 100\% \\
 &= \frac{32+15}{32+15+3+0} \times 100\% \\
 &= \frac{47}{50} \times 100\% \\
 &= 94\%
 \end{aligned}$$

Hasil AUC (Area Under Curve) dari pengujian pertama :



Sumber Penelitian 2019

Gambar 7.
Kurva AUC Pengujian 1

b. Pengujian K-Fold Validation⁶

Pengujian ini dilakukan dengan membagi sebanyak 6 bagian pada 50 data kuisioner yang akan diuji. 6 bagian tersebut terdiri dari 5 bagian data

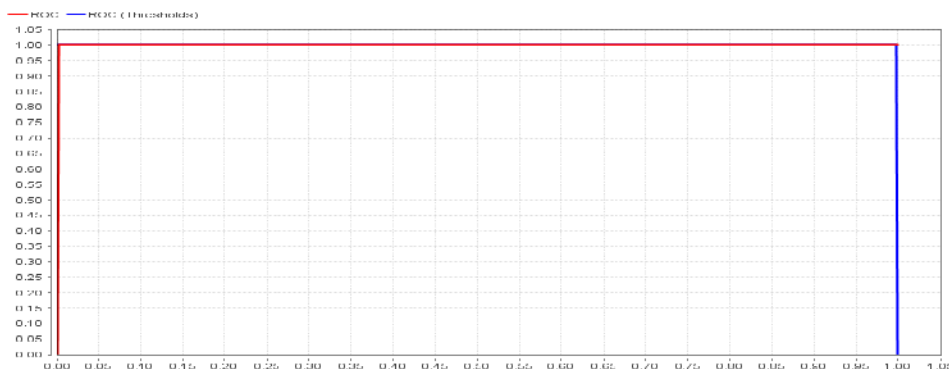
Tabel 3.

Tabel Confusion Matrix Pengujian 2

	Actual	
	Puas	Tidak Puas
Pred. Puas	35	0
Pred. Tidak Puas	0	15

Jumlah true positive (tp) 35 data, false positive (fp) 0 data, true negative (tn) 15 data, dan false negative 0 data. Dari hasil pengujian 2 pada Rapid Miner diperoleh akurasi sebesar 100%. Berikut adalah perhitungan akurasi:

AUC: 0.500 +/- 0.000 (micro average: 0.500) (positive class: Tidak Puas)



Gambar 8.

Kurva AUC Pengujian 2

c. Pengujian K-Fold Validation⁴

Pengujian ini dilakukan dengan membagi sebanyak 4 bagian pada 50 data kuisioner yang akan diuji. 4 bagian tersebut terdiri dari 3 bagian data training dan 1 bagian data testing. Dan dalam pengujian ini akan menghasilkan nilai akurasi, precision, recall dan AUC (Area Under Curve). Hasil akurasi confusion matrix pada pengujian 3 :

Tabel 4

Tabel Confusion Matrix pengujian 3

	Actual	
	Puas	Tidak Puas
Pred. Puas	33	0
Pred. Tidak Puas	2	15

training dan 1 bagian data testing. Dan dalam pengujian ini akan menghasilkan nilai akurasi, precision, recall dan AUC (Area Under Curve).

Hasil akurasi confusion matrix pada pengujian 2:

$$\begin{aligned}
 \text{Akurasi} &= \frac{tp+tn}{tp+tn+fp+fn} \times 100\% \\
 &= \frac{35+15}{35+15+0+0} \times 100\% \\
 &= \frac{50}{50} \times 100\% \\
 &= 100\%
 \end{aligned}$$

Hasil AUC (Area Under Curve) dari pengujian kedua :

Jumlah true positive (tp) 33 data, false positive (fp) 2 data, true negative (tn) 15 data, dan false negative 0 data. Dari hasil pengujian 2 pada Rapid Miner diperoleh akurasi sebesar 96%. Berikut adalah perhitungan akurasi:

$$\begin{aligned}
 \text{Akurasi} &= \frac{tp+tn}{tp+tn+fp+fn} \times 100\% \\
 &= \frac{33+15}{33+15+2+0} \times 100\% \\
 &= \frac{48}{50} \times 100\% \\
 &= 96\%
 \end{aligned}$$

Hasil AUC (Area Under Curve) dari pengujian ketiga :



Sumber Penelitian 2019

Gambar 9.
Kurva AUC Pengujian 3

Dari tiga kali pengujian dengan nilai *K-Fold Validation* yang berbeda menghasilkan perbandingan seperti pada tabel berikut :

Tabel 5.
Tabel Perbandingan Akurasi

<i>K-Fold Validation</i>	Akurasi	AUC
10	94%	0.500
6	100%	0.500
4	96%	0.597

4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan bahwa: a) Kepuasan pengunjung Setu Babakan dapat diprediksi dan dievaluasi dengan memanfaatkan teknik data mining menggunakan algoritma *decision tree* C4.5 untuk mengetahui variabel apa saja yang mempengaruhi kepuasan pengunjung dengan data training yang telah diperoleh. b) Dari 3 kali percobaan dan pengujian prediksi kepuasan pengunjung Setu Babakan dengan *rapidMiner* menggunakan metode *decision tree* C4.5, diperoleh akurasi sebesar 100% dengan menggunakan *confusion matrix*.

5. REFERENSI

- [1] Takalapeta, S. (2018). Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode C4.5. *Jurnal Informatika Merdeka Pasuruan*, 3(3), 34–38.
- [2] Nofriansyah, D., & Nurcahyo, G. W. (2015). *Algoritma Data Mining Dan Pengujian*. Deepublish.
- [3] Shiddiq, A., Kumalasari Niswatin, R., & Farida, I. N. (2018). Analisa Kepuasan Konsumen Menggunakan Klasifikasi

Decision Tree Di Restoran Dapur Solo (Cabang Kediri). *Generation Journal*, 2(1), 9–18.

- [4] Rangkuti, F. (2013). *Customer Service Satisfaction & Call Center Berdasarkan ISO 9001*. Gramedia Pustaka Utama.