Journal of Artificial Intelligence and Engineering Applications
Website: https://ioinformatic.org/

15th October 2021. Vol. 1. No. 1

Application of the C5.0 Algorithm to Determine the Level of Public Satisfaction with the E-KTP Recording Service at the Bandar Sub-District Office

Dini Fadila Hardani[1]*, Poningsih[2], Yuegilion Pranayama Purba[3]

[1,3]STIKOM Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia
[2]AMIK Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia
[1,2,3]Jln. Sudirman Blok A No. 1-3 Pematangsiantar, Sumatera Utara
[1]dinifadilahardani15@gmail.com, [2]poningsih@amiktunasbangsa.ac.id

Abstract

Community satisfaction at the Bandar Sub-district Office is one of the most important things in assessing the level of e-KTP recording services provided by the agency to the community. The purpose of this study was to determine the quality of the e-KTP recording service at the Bandar Sub-district Office in terms of the Service Procedure, Time, Behavior and Facilities aspects of the Bandar sub-district community. At the Bandar Camat Office these four aspects have not been measured with certainty, so the agency finds it difficult to determine which aspects must be improved. The method used in this study is the C5.0 Algorithm, where the data source used is a questionnaire/questionnaire technique given to the people of Bandar sub-district. The research test process uses Rapid Miner software to create a decision tree. The results of the study obtained 12 rules for classifying the level of community satisfaction with e-KTP recording services. The C5.0 algorithm can be used in cases of community satisfaction with an accuracy rate of 100%. From these results, it is expected to improve the quality of service for the e-KTP recording of the Bandar Sub-District Office to be even better.

Keywords: *Data Mining, Society, C5.0 Algorithm, RapidMiner*

1. Introduction

Community satisfaction is considered as one of the most important quality benchmarks and one of the main indicators of government facilities [1]. This is an e-KTP recording service provided by the agency and this is an important component of the reason for measuring people's satisfaction [2][3]. According to Sugianto and Apandi, the implementation of this public service does not provide the best results if government agencies cannot operate in the best way, it must be balanced with optimizing the performance of government officials to do so consistently by paying attention to all the needs and expectations of the people [4]. In the process of providing community services, it is very important, basically service is a direct interaction activity between one person and another so as to provide quality service [5]. Service quality refers to the extent to which resources are used fairly, efficiently and effectively within the limited capacity of government and society with good professional standards to meet community needs and to be carried out safely and satisfactorily [2], [6]. In reality, it can be seen that there are still many complaints from the public regarding satisfaction services for e-KTP recordings carried out by government officials, which many find services that have not been maximized and the quality of human resources is not adequate. Service problems are not difficult and complicated, but if this is not considered, it can cause problems for the agency [7].

In this study, data sampling was carried out at random and did not focus on one person in the community [8], [9]. The data used in this study are primary data. The sample data used in this study were 100 respondents using 4 variables, including Service Procedures, Time, Behavior and Facilities. From the problems above, the authors use classification data mining techniques with the C5.0 Algorithm to get a more efficient solution to community satisfaction with e-KTP recording services. The C5.0 algorithm is a decision tree-based algorithm which is a refinement of the ID3 and C4.5 algorithms formed by Ross Quinlan in 1987 [10], [11], [12]. The C5.0 algorithm can be used to examine various things, including research related to predicting cleaning service acceptance, the results obtained from the research can be seen from the criteria of the employment contract [10]. With the classification of public satisfaction with e-KTP recording services, efforts can be made to improve service quality [9], [13].

2. Results and Discussion

In this study, the authors obtained data from the questionnaire results from the Bandar Camat office with a total of 100 respondents. To make as a root, based on the highest gain value of the existing attribute. To calculate the gain, use a formula like the following [5][14][15]:

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

Information:

A = Attribute

S = Set of cases

S1 = Number of samples

Before getting the gain value, the entropy value is sought first. To find the value of entropy use the following formula:

$$E\ (s) = \sum_{i=1}^{n} - pi * log_2\ (Pi) \quad (2)$$

Information :

S = case set

S1 = number of samples

Pi = proportion

2.1. Pengolahan Data Menggunakan Algoritma C5.0

The following is the data obtained from the community satisfaction survey of the Bandar Camat office:

Table 1. Questionnaire Data from the Bandar Sub-District Office

Respondent	Procedure	Time	Behavior	Fasility	Results
R1	Easy	Fast	Polite and friendly	Enough	Satisfied
R2	Easy	Very Fast	Polite and friendly	Very good	Satisfied
R3	Easy	Fast	Polite and friendly	Good	Satisfied
R4	Easy	Fast	Polite and friendly	Good	Satisfied
R5	Easy	Fast	Polite and friendly	Good	Satisfied
R6	Very Easy	Fast	Very Polite and friendly	Good	Satisfied
R7	Easy	Fast	Polite and friendly	Good	Satisfied
R8	Less Easy	Not Fast	Polite and friendly	Enough	Not satisfied
R9	Easy	Fast	Polite and friendly	Good	Satisfied
R10	Easy	Not Fast	Polite and friendly	Enough	Satisfied
R11	Easy	Not Fast	Polite and friendly	Good	Satisfied
R12	Easy	Not Fast	Polite and friendly	Good	Satisfied
R13	Easy	Fast	Polite and friendly	Enough	Satisfied
R14	Not Easy	Not Fast	Polite and friendly	Enough	Not satisfied
R15	Less Easy	Not Fast	Polite and friendly	Good	Satisfied
R16	Less Easy	Not Fast	Less Polite and friendly	Enough	Not satisfied
R17	Easy	Fast	Polite and friendly	Good	Satisfied
R18	Easy	Not Fast	Less Polite and friendly	Enough	Not satisfied
R19	Not Easy	Not Fast	Not Polite and friendly	Not good	Not satisfied
R20	Less Easy	Fast	Polite and friendly	Good	Satisfied
R21	Very Easy	Fast	Polite and friendly	Good	Satisfied
R22	Very Easy	Fast	Very Polite and friendly	Good	Satisfied
R23	Not Easy	Not Fast	Not Polite and friendly	Not good	Not satisfied
R24	Easy	Fast	Polite and friendly	Good	Satisfied
…..	…..	…..	…..	…..	…..
R95	Easy	Not Fast	Less Polite and friendly	Enough	Not satisfied
R96	Easy	Fast	Less Polite and friendly	Good	Satisfied
R97	Easy	Fast	Polite and friendly	Good	Satisfied
R98	Easy	Not Fast	Polite and friendly	Enough	Satisfied
R99	Easy	Fast	Polite and friendly	Good	Satisfied
R100	Easy	Fast	Less Polite and friendly	Good	Satisfied

The calculation of the C5.0 algorithm begins by selecting the root attribute first by finding the total number of cases, the number of decision cases is not difficult. Calculating the entrophy of all cases divided based on Service Procedures, Time, Behavior and Facilities. After that, the gain is calculated for each attribute. The calculation results are shown in the following table.

Table 2 Calculation of Node 1

Node 1		Number of Cases	Satisfied	Not satisfied	Entropy	Information Gain
Total		100	70	30	0,8812909	
Service Procedure					100	0,108250936
	Very Easy	6	6	0	0	

	Easy	79	59	20	0,816255409	
	Less Easy	9	4	5	0,99107606	
	Not Easy	6	1	5	0,650022422	
Time						0,453465222
	Very Fast	2	2	0	0	
	Fast	52	52	0	0	
	Not Fast	36	13	23	0,943601631	
	Not Fast	10	3	7	0,881290899	
Behavior						0,36049466
	SSDR	3	3	0	0	
	SDR	60	56	4	0,353359335	
	KSDR	34	11	23	0,908178347	
	TSDR	3	0	3	0	
Facility						0,210882891
	Very good	5	4	1	0,721928095	
	Good	36	33	3	0,41381685	
	Enough	45	31	14	0,894451885	
	Bad	14	2	12	0,591672779	

From the results of calculations in table 2, it is obtained that the attribute that becomes the node (root) is the time that has the highest gain, namely 0.453465222, which consists of 4 sub-attributes, namely Very Fast, Fast, Not Fast and Not Fast. To determine the next root node, a decision tree can be drawn from the table above as follows:

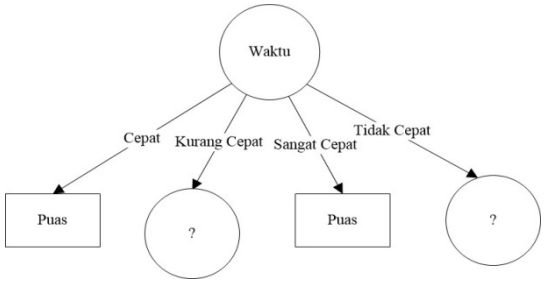

Figure 1. Decision Tree 1

In looking for the results of the next calculation on the root node Time - Not Fast until the final result is obtained at node 1.2. can be seen in the following table:

Table 3. Calculation of Node 1.2

Node 1.2		Number of Cases	Satisfied	Not satisfied	Entropy	Information Gain
Time-KR		10	3	7	0,8812909	
Service Procedure	Very Easy	0	0	0	0	0,796388649
	Easy	4	1	3	0,8112781	
	Less Easy	2	1	1	1	
	Not Easy	4	1	3	0,8112781	
Behaviour						0,991290899
	SSDR	0	0	0	0	
	SDR	3	3	0	0	
	KSDR	4	0	4	0	
	TSDR	4	0	4	0	
Facility						0,811521423
	Very good	3	2	1	0,9182958	
	Good	3	0	3	0	
	Enough	2	1	1	1	
	Bad	2	0	2	0	

From the results of the calculations in table 6, the attribute that becomes the time branch node - Not Fast where the behavior has obtained a decision so that the calculation is complete, the decision tree can be described from the table above as follows:

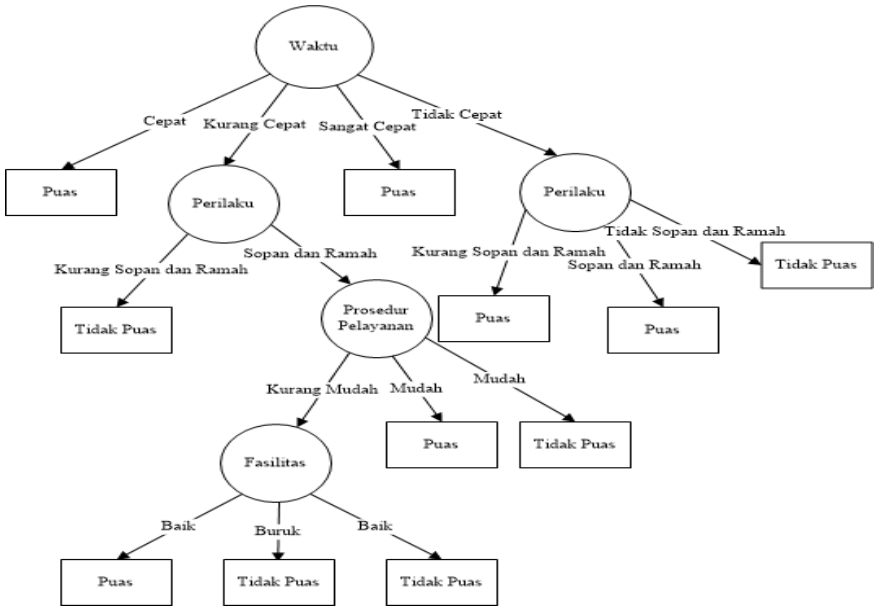

Figure 5. Final Node Decision Tree

2.2. Testing With RapidMiner

Testing the data with Rapidminer produces the same rules as the manual calculation of the C5.0 Algorithm as shown in Figure 6 below:

Tree

```
Waktu = Cepat: Puas {Puas=52, Tidak Puas=0}
Waktu = Kurang Cepat
|   Perilaku = Kurang Sopan dan Ramah: Tidak Puas {Puas=0, Tidak Puas=19}
|   Perilaku = Sopan dan Ramah
|   |   Prosedur Pelayanan = Kurang Mudah
|   |   |   Fasilitas = Baik: Puas {Puas=1, Tidak Puas=0}
|   |   |   Fasilitas = Buruk: Tidak Puas {Puas=0, Tidak Puas=1}
|   |   |   Fasilitas = Cukup Baik: Tidak Puas {Puas=0, Tidak Puas=1}
|   |   Prosedur Pelayanan = Mudah: Puas {Puas=12, Tidak Puas=0}
|   |   Prosedur Pelayanan = Tidak Mudah: Tidak Puas {Puas=0, Tidak Puas=2}
Waktu = Sangat Cepat: Puas {Puas=2, Tidak Puas=0}
Waktu = Tidak Cepat
|   Perilaku = Kurang Sopan dan Ramah: Tidak Puas {Puas=0, Tidak Puas=4}
|   Perilaku = Sopan dan Ramah: Puas {Puas=3, Tidak Puas=0}
|   Perilaku = Tidak Sopan dan Ramah: Tidak Puas {Puas=0, Tidak Puas=3}
```

Figure 6. Description of the Decision Tree

The picture above shows the results of a complete description of the decision tree that has been formed using the C5.0 algorithm. the results of the description also show that the use of data mining algorithm C5.0 Good is used in the process of extracting data to draw some conclusions which are visualized with a decision tree. The following rules are generated from the decision tree:

Table 7. Generated rules

No.	Rules	Decision
1.	IF Time = Fast	Satisfied
2.	IF Time = Not Fast dan Behavior = Less Polite and friendly	Not satisfied
3.	IF Time = Not Fast dan Behavior = Polite and friendly dan Service Procedure= Less Easy dan Facility = Good	Satisfied
4.	IF Time = Not Fast dan Behavior = Polite and friendly dan Service Procedure= Less Easy dan Facility = Buruk	Not satisfied
5.	IF Time = Not Fast dan Behavior = Polite and friendly dan Service Procedure= Less Easy dan Facility = Enough	Not satisfied

6.	IF Time = Fast dan Service Procedure= Easy	Satisfied
7.	IF Time = Fast dan Service Procedure= Not Easy	Not satisfied
8.	IF Time = Very Fast	Satisfied
9.	IF Time = Not Fast dan Behavior = Less Polite and friendly	Not satisfied
10.	IF Time = Not Fast dan Behavior = Polite and friendly	Satisfied
11.	IF Time = Not Fast dan Behavior = Not Polite and friendly	Not satisfied

3. Conclusion

Based on all the results of the research stages that have been carried out on the application of the C5.0 Algorithm for satisfication level classification and e-KTP recording services, it can be concluded that Satisfaction and the public for e-KTP recording services can apply the C5.0 algorithm. There are 4 attributes that support the level of satisfaction and the community used is the Service Procedure, Time, Behavior and Facility and the determination obtained in the study is Satisfied with the resulting accuracy level produced by this method is 100%.

References

[1] G. R. P, A. P. Windarto, E. Irawan, W. Saputra, and H. Okrapana, "Penerapan Data Mining Menggunakan Algoritma C4 . 5 Dalam Mengukur Tingkat KeSatisfiedan Pasien BPJS," *Pros. Semin. Ris. dan Inf. sains*, vol. 2, pp. 376–385, 2020.

[2] S. Suleman, "Kualitas Pelayanan E-KTP di Dinas Kependudukan dan Catatan Sipil Kabupaten Halmahera Selatan," *J. Adm. Publik*, vol. Vol.5, no. No.1, p. 3, 2019, [Online]. Available: http://journal.unismuh.ac.id/index.php/kolaborasi.

[3] R. Ester, A. Sihombing, J. T. Hardinata, Z. Masruro, S. T. Bangsa, and S. Utara, "Penerapan Algoritma C4 . 5 Data Mining Dalam Mengukur Tingkat KeSatisfiedaan Masyarakat Kecamatan Siantar Terhadap PerGoodan Jalan 2019," vol. 2, no. 2, pp. 91–98, 2019.

[4] T. H. Apandi, C. A. Sugianto, and C. R. Service, "Algoritma Naive Bayes untuk Prediksi KeSatisfiedan Pelayanan Perekaman e-KTP (Naive Bayes Algorithm for Satisfaction Prediction of e-ID," vol. 7, no. November, pp. 125–128, 2019.

[5] E. S. Pribadi, P. Poningsih, and H. S. Tambunan, "Analisa Tingkat KeSatisfiedan Masyarakat Terhadap Pelayanan Pengadilan Agama Pematangsiantar Menggunakan Algoritma C4.5," *Brahmana J. Penerapan Kecerdasan Buatan*, vol. 2, no. 1, pp. 33–40, 2020, doi: 10.30645/brahmana.v2i1.46.

[6] R. S. Septarini *et al.*, "Rancang Bangun Aplikasi Customer ' Ku Untuk Menentukan Kualitas Pelayanan Menggunakan Metode Simple Additive," pp. 552–559, 2018.

[7] F. N. Widowati and M. Wahyudin, "Pengaruh Citra Merek, Kualitas Pelayanan dan Kualitas Produk Terhadap Loyalitas Pelanggan (Studi pada Konsumen Toko Javascloth Mojolaban Sukoharjo)," 2018, [Online]. Available: http://eprints.ums.ac.id/id/eprint/67658.

[8] I. A. Nikmatun and I. Waspada, "Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *J. SIMETRIS*, vol. 10, no. 2, pp. 421–432, 2019.

[9] H. Dhika, F. Destiawati, and A. Fitriansyah, "Implementasi Algoritma C4. 5 terhadap KeSatisfiedan Pelanggan," pp. 80–86, 2018, doi: 10.31227/osf.io/fgc7a.

[10] R. P. Padang, "IMPLEMENTASI DATA MINING ALGORITMA C5 . 0 DALAM MEMPREDIKSI PENERIMAAN CLEANING SERVICE (CS) PADA PT ISS INDONESIA MEDAN," vol. 6, pp. 304–309, 2019.

[11] Carlis Hutabarat, "Penerapan Data Mining Untuk Memprediksi Permintaan Produk Kartu Perdana Internet Menggunakan Algoritma C5.0 (Studi Kasus: Vidha Ponsel)," *Pelita Inform.*, vol. 6, no. April, pp. 419–424, 2018.

[12] S. W. Sinaga, "Darma Medan Berdasarkan Daerah Dengan Menggunakan Algoritma," vol. 18, no. April, pp. 226–234, 2019.

[13] N. M. Wasilah, *Penerapan Data Mining Untuk Menganalisis KeSatisfiedan Konsumen Menggunakan Metode C4.5*, vol. 5. 2013.

[14] T. Permana, A. M. Siregar, A. F. N. Masruriyah, and A. R. Juwita, "Perbandingan Hasil Prediksi Kredit Macet Pada Koperasi," *Conf. Innov. Appl. Sci. Technol.*, vol. 3, no. 1, pp. 737–746, 2020, [Online]. Available: http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/article/view/1970.

[15] N. Jenggawah, S. Pada, K. Berpikir, K. Dan, and M. Belajar, *Digital Digital Repository Repository Universitas Universitas Jember Jember Digital Jember Digital Repository Repository Universitas Universitas Jember*. 2010.