

ANALYZING CUSTOMER SATISFACTION ON ARTHASPA SERVICES WITH THE C4.5 ALGORITHM APPROACH

Sofian Wira Hadi^{1*}, Wawan Kurniawan², Kudiantoro Widiyanto³, Ibnu Alfarobi⁴

^{1,2}Computer science

^{1,2}STMIK Nusa Mandiri

www.nusamandiri.ac.id

¹14002361@nusamandiri.ac.id ; ²14002315@nusamandiri.ac.id ;

^{3,4}Accounting Information Systems

^{3,4}Bina Sarana Informatika University

www.bsi.ac.id

³kudiantoro.kdw@bsi.ac.id; ⁴ibnu.iba@bsi.ac.id

(*) Corresponding Author

Abstract

Customer or buyer satisfaction is closely related to how a seller of services or a store serves its visitors. Good service for visitors also makes a good impression on visitors, while the opposite will cause a very bad or unfavourable impression in the eyes of customers, and may also lead to the reluctance of visitors to come back and lose the seller's opportunity to get potential buyers to become customers. This study attempts to analyze customer satisfaction with the services provided by Arthaspa outlets in Grand Kemang hotels using the C4.5 Algorithm approach. The attributes used are comfort, cleanliness, tidiness, and price. samples taken are customers who have transacted at least once.

Keywords: Data Mining, C4.5, Classification, Arthaspa

Abstrak

Kepuasan pelanggan atau pembeli setuju dengan layanan penjual atau penyedia layanan yang melayani pengunjungnya. Layanan yang baik untuk pengunjung yang menarik perhatian kedua pengunjung, sementara sebaliknya akan menarik perhatian yang lebih baik atau kurang di mata pelanggan, dan mungkin juga menarik pengunjung yang datang kembali untuk membeli kesempatan membeli pembeli yang berpotensi mendapatkan pelanggan. Penelitian ini mencoba menganalisis kepuasan pelanggan dengan layanan yang disediakan oleh outlet Arthaspa yang tersedia di Grand Kemang Hotel menggunakan Algoritma C4.5. Atribut yang digunakan adalah kenyamanan, kebersihan, kerapian, dan harga. Sampel yang diambil adalah pelanggan yang pernah bertransaksi sekali. Studi ini mendapatkan kesimpulan paling populer tentang simpul "harga" saat menggunakan Microsoft Excel dan "Kebersihan" saat menggunakan perangkat lunak Rapidminer.

Kata Kunci: Data Mining, C4.5, Klasifikasi, Arthaspa.

INTRODUCTION

The best quality of service to customers will also affect the trust of customers in the company so that customers are satisfied with the services received and thus the customer will convey satisfaction to others, this case makes the market share expand and the company will be superior to its competitors (AFRIZAL, 2018).

Consumer satisfaction is a comparison between the services received and consumer expectations, consumers have assessed satisfaction or dissatisfaction with the level of expectations they create in mind. A situation of

dissatisfaction occurs if consumers after using the product or service purchased feel that performance does not match consumer expectations (Alawiyah, 2018).

Quality of service is very important and closely related to customer satisfaction itself. Good service quality will give satisfaction from customers, while service quality is lacking or not good it has an uncomfortable effect on the customer (Sobandi, 2019) the data visit, and it is also possible that over time can cause customers to switch to other competitors (Hartono, 2017) who has a similar business. The current tight competition also forces the seller or buyer to

produce a good service so that it can satisfy the customer (Budiarti, 2018).

Quality of service is very important and closely related to customer satisfaction. A good quality of service will provide satisfaction from customers, while a quality of service that is lacking or not good will have an uncomfortable effect for consumers who come to visit, and it is also possible that over time it can cause customers to switch to other competitors who have similar businesses. The current competitive conditions also force sellers or buyers to produce good service so that they can satisfy customers. Companies in knowledge of the market or customers have a positive correlation in sales performance (Noyita, 2019).

Arthaspa also prioritizes customer satisfaction in marketing strategies for the advancement of its company because it is related to market share and customer retention (Puung et al., 2014). Besides arthaspa also prioritizes good performance by employees so that the assessment produced by consumers of the service is also good and it is very important for the company, because customer satisfaction is a pleasant response and can be fulfilled, while customer dissatisfaction is an unexpected disappointment (Susi et al., Nd).

In this research, it is expected that the results of customer satisfaction will be analyzed using the data mining approach C4.5 because the C4.5 algorithm is easy to understand and interpreted in its use (Eki, 2016). To find out whether Artspa has provided services that are in accordance with customer wishes. So the evaluation needs to be done from the customer side considering the attributes of comfort, cleanliness, tidiness, and price are very important in providing satisfaction to customers.

RESEARCH METHODOLOGY

Knowledge Discovery in Database (KDD)

Data Mining or Knowledge Discovery in Database (KDD). The KDD process is the result of minimal data extracting a data pattern, and altering the results so that they are easy to understand (Riandari & Simangunsong, 2019). In KDD there are six most basic elements in KDD information retrieval techniques, namely:

1. Working on data that will be processed with many sources.
2. Efficient use of data is required
3. Prioritizing statutes
4. Requires a high level of language usage
5. Use several forms of automatic learning
6. Produce unique results

The Knowledge Discovery in Database (KDD) process can be outlined as follows (Yunita, 2018) :

1. Data Selection is the process of selecting words from a collection of data, data selection is done before the stage of obtaining information in KDD. the results of the data selection will be stored in a file, separate from the database Pre-processing or Cleaning is before the data mining process is carried out, it is necessary to clean up the data that is the focus of KDD. The cleaning process includes removing duplicate data, checking for inconsistent data, and correcting data errors;
2. Transformation is performed on data that has been selected or selected. so that the data selected is in accordance with the provisions of the data mining process.
3. Data mining is the science of finding interesting patterns or information in large amounts of data by using certain techniques or methods. techniques, methods or algorithms in data mining vary widely. the choice of data mining method is very dependent on the KDD process (Mardi, 2017).
4. Interpretation or evaluation resulting from the data mining process really needs to be defined in a form that is easy to understand by interested parties. this stage is one part of the KDD process called interpretation. this stage as an evaluation of whether the pattern or information found is contrary to the previous hypothesis.

Data Mining

Data mining is the science of extracting information from big data, in accordance with the purpose of data mining, which is to make a decision from a large data and stored in a database, data warehouse or information stored from a repository (Tarigan et al., 2017).

Data collection

Data collection method is done by observation directly and use the Questionnaire to get accurate data.

1. Observation is done by collecting sales data by visiting directly the hotel in collaboration with CV. Artha Gemilang to get the information needed related to research.
2. Questionnaire where the researcher distributes a list of questions to respondents regarding various aspects related to the value of each customer satisfaction attribute which will later be carried out on the results of research and discussion.

Research Population

The population that will be used is the total number of customers at outlets that work with CV, Artha Spa since the beginning of January 2019 until the end of June 2019 who have done at least once. ie as many as 186 people.

Research Sample

Determination of the number of study samples is calculated by the Slovin formula (Luvia et al., 2017):

$$n = \frac{N}{1 + Ne^2} \dots\dots\dots (1)$$

Where:

n: number of samples

N: total population

E: error tolerance limit of 5 percent or 0.05.

$$\text{Number of respondents to use: } \frac{200}{1 + (200 * 0,0025)}$$

Yields n = 133.33, if rounded to 133. So the number of respondents used is 127 people.

Data Analysis Method

The steps taken in data analysis include:

- 1 The results of the respondents' answers in the questionnaire were converted into a Likert scale
- 2 Obtained Likert Scale then to make a decision tree with the C4.5 algorithm approach. Begin by forming which attribute will become root, or which root attribute will be based on the highest gain gain. If the value of an attribute has not resulted in a unanimous decision, then a recalculation is made by making a new branch under the previous node, but if a unanimous decision has occurred, then the calculation will be stopped and a final conclusion obtained.
- 3 The results of the algorithm calculation are then represented as a decision tree shape.

RESEARCH RESULTS AND DISCUSSION

As for the research on Arthaspa using premiere data taken from May to June 2019 for the results of the questionnaire can be seen in Table 1:

Table 1. Questionnaire Results

No	Attribute	Number of cases	Yes	Not
	Total	50	32	18
1	Convenience			
	Very comfortable	31	27	4
	Enough	10	5	5
	Less comfortable	9	0	9

No	Attribute	Number of cases	Yes	Not
	Total	50	32	18
2	Cleanliness			
	Clean	41	32	9
	Not clean	9	9	9
3	Neatness			
	Neat	29	23	6
	Not neatly	21	9	12
4	Price			
	Affordable	22	22	0
	Relatively inexpensive	24	7	17
	Expensive	4	3	1

Evaluation and Validation

Evaluation and validation are the results of the classification of data that has been determined based on the process used, henceforth after knowing the evaluation of the classification model based on the number of dataset records that have been predicted correctly and incorrectly in the classification modeling, these results can be known as confusion matrix. after getting a number of attributes, the next step is processing the selection of attributes(Santoso, 2014). This attribute selection is done to get the attributes whose values are relevant. the following explanation of the attributes used:

1. Convenience is an attribute given to respondents to assess comfort in service and is grouped into 3 categories, namely, Very comfortable, quite, less comfortable.
2. Cleanliness is an attribute to assess the cleanliness of the environment in the artha spa, and is grouped into 2 categories, clean and unclean.
3. Neatness is an attribute that assesses the neatness of arthaspa employees who are grouped into 2 categories, neat and untidy.
4. Price is an attribute of the price offered by artha spas to consumers, and is categorized into 3, namely affordable, relatively inexpensive and expensive.

In this research the test results can be seen from the following steps:

Information Gain and Entrhopy

The first step to do is calculate the entropy value and information gain data in Table 2:

Table 2. Entrhopy and Gain Values for all attributes

Attribute	Amount (s)	Yes (Si)	No (Si)	Entrhopy	Gain
Total	50	32	18	0.942683	
Convenience					0.398721
Very	31	27	4	0.5547	



Attribute	Amount (s)	Yes (Si)	No (Si)	Entropy	Gain
comfortable				78	
Enough	10	5	5	1	
Less comfortable	9	0	9	0	
Cleanliness					0.320077
Clean	41	32	9	0.759276	
Not clean	9	9	9	0	
Neatness					0.102292
Neat	29	23	6	0.735509	
Not neatly	21	9	12	0.985228	
Price					0.459766
Affordable	22	22	0	0	
Relatively inexpensive	24	7	17	0.870864	
Expensive	4	3	1	0.811278	

Table 3. Price Attributes

Attribute	Amount (s)	Yes (Si)	No (Si)	Entropy	Gain
Total	50	32	18	0.942683	
Price					0.459766
Affordable	22	22	0	0	
Relatively inexpensive	24	7	17	0.870864	
Expensive	4	3	1	0.811278	

In Table 3 the Price Attributes have 3 categories, namely, affordable, relatively cheap, expensive. With each entraping value. because the affordable category has a value of 0, what we are looking for in the next node is relatively cheap and expensive, it can be seen in Table 4:

Table 4. Relatively Cheap

Attribute	Amount (s)	Yes (Si)	No (Si)	Entropy	Gain
Total	50	32	18	0.942683	
Convenience					0.766425
Very comfortable	10	7	3	0.881291	
Enough	5	0	5	0	
Less comfortable	9	0	9	0	
Cleanliness					0.626299
Clean	16	7	9	0.988699	
Not clean	8	0	8	0	
Neatness					0.581891
Neat	5	0	5	0	

Attribute	Amount (s)	Yes (Si)	No (Si)	Entropy	Gain
Not neatly	19	7	12	0.949452	

The results of searching for the relatively inexpensive category price attribute yielded the highest Gain value, namely comfort with a value of 0.766425. For further searching for the gain value for the convenience attribute, can be seen in Table 5:

Table 5. Comfort

Attribute	Qty (s)	Yes (Si)	No (Si)	Entropy	Gain
Total	50	32	18	0.942683	
Cleanliness					0.942683
Clean	7	7	0	0	
Not clean	3	0	3	0	
Neatness					0.766425
Neat	0	0	5	0	
Not neatly	10	7	3	0.881291	

The results of looking for comfort attributes produced the highest Gain value, namely cleanliness with a value of 0.942683. Because the entropy value of the comfort category 0 for the next calculation is to find the gain value from the expensive category, it can be seen in Table 6:

Table 6. Expensive Categories

Total Attribute	number (s)	Yes (Si)	No (Si)	Entropy	Gain
	50	32	18	0.942683	
Convenience					0.877781
Very comfortable	4	3	1	0.811278	
Enough	0	0	0	0	
Less comfortable	0	0	0	0	
Cleanliness					0.942683
Clean	3	3	0	0	
Not clean	0	0	0	0	
Neatness					0.877781
Neat	4	3	1	0.811278	
Not neatly	0	0	0	0	

Search results from the expensive category produced the highest Gain value, namely cleanliness with a value of 0.942683. Because the entropy value of the comfort category 0 for the node calculation is complete.

Decision Tree



Figure 1. Decision Tree Results

The results of the decision tree modeling in Figure 1 shows that all cases are included in the class determined and therefore the decision tree is the final decision tree. after the results of modeling with the decision tree the creation of rules in the decision tree. the rules that can be formulated are as follows:

"If Cleanliness = not clean then class = Not satisfied"

"If Cleanliness = Clean and comfort = very comfortable, class = satisfied"

"If Cleanliness = Clean and comfort = Less comfortable then class = not satisfied"

"If Cleanliness = Clean and comfort = enough and price = relatively cheap, class = not satisfied"

"If Cleanliness = Clean and comfort = enough and price = affordable, class = satisfied"

The results obtained in the sample data that are rooted are the service attributes in the decision tree, while attributes such as comfort, cleanliness and price are good food. from the sample data, the number of rules used formed 5 rules.

CONCLUSIONS AND SUGGESTIONS

Conclusion

From the results obtained in the previous discussion, conclusions can be drawn, with the attributes used such as: Comfort, cleanliness, tidiness and price greatly affect the level of customer satisfaction. If cleanliness is not clean then the customer will be dissatisfied if the cleanliness is clean and very comfortable then the customer will be satisfied but if the cleanliness of the comfort level is sufficient and the price is affordable then the customer will be satisfied with the existing services.

Suggestion

It needs to be evaluated and recalculated regularly so that the company can continue to optimize its services to customers. It is recommended to do another algorithm approach, included in the classification category and seen the highest level of accuracy in order to be taken into consideration for analyzing the next level of customer satisfaction. And of course it can also be a constructive suggestion

REFERENCE LIST

- Afrizal, B. (2018). The Effect of Service Quality and Price on Customer Satisfaction at Warung Mbah Ganis. Master of Management, Nusantara Business Institute, 1, 10-18.
- Alawiyah, T. (2018). The Influence of Service Quality on Consumer Satisfaction Staying at Diamond Hotels in Samarinda City. *eJournal of Business Administration*, 6 (2), 803–817.
- Budiarti, A. (2018). The Effect of Service Quality and Complaint Handling on Customer Satisfaction and Loyalty of Sharia Commercial Banks in Anindhya Budiarti. *Equity*, 15 (110), 210-231.
- Eki, RM (2016). Customer Satisfaction Analysis Using C4 Algorithm. 5. Faculty of Computer Science, 3 (2), 1–14.
- Hartono, A. (2017). Analysis Of Service Quality Influence On Customer Satisfaction In Forming Customer Loyalty In The Naked Crab Restaurant Surabaya. *Journal of Hospitality and Services Management*, 5 (2), 174–187.
- Luvia, YS, Windarto, AP, Solikhun, S., & Hartama, D. (2017). Application Of C4.5 Algorithm For Classification Of Student Success Predictions In Amik Tunas Nation. *Jurasik (Journal of Information Systems and Informatics Engineering Research)*, 1 (1), 75. <https://doi.org/10.30645/jurasik.v1i1.12>
- Mardi, Y. (2017). Data Mining: Classification Using C4.5 Algorithm. *Journal of Edict Informatics*, 2 (2), 213-219.
- Noyita, E. (2019). Measuring Improvement Of Service Quality And Customer Trust On Customer Satisfaction With Sem Analysis

- Method (Study at Kampoeng Spa Batam. Khazanah Ilmu Berazam, 2, 62-70.
- Puung, FK, Fudholi, A., & Dharmmesta, BS (2014). Analysis of the Effect of Service Quality on Customer Satisfaction and Loyalty in Salons and Spas. Analysis of the Effect of Service Quality on Customer Satisfaction and Loyalty in Salons and Spas, 4 (2), 105-110. <https://doi.org/10.22146/jmpf.274>
- Riandari, F., & Simangunsong, A. (2019). Application of C4.5 Algorithm to Measure Student Satisfaction Level. DIKTI Accredited, 3 (2), 1-7.
- Santoso, TB (2014). Analysis and Application of C4.5 Method For Predicting Customer Loyalty. LIMIS Faculty of Engineering Scientific Journal, 10 (1), 1-6.
- Sobandi, MM (2019). SERVICE MANAGEMENT IN PHARMACIES P. BHAKTI KENCANA TOWARDS CUSTOMER SATISFACTION. Journal of Pharmacopolium, 2 (1). <https://doi.org/10.36465/jop.v2i1.468>
- Susi, I., Situmorang, S., Nasari, F., Systems, J., University, I., Utama, P., & Utama, UP (nd). Analysis of customer satisfaction with spa services with servqual (case study: pt. Royal amadeus). December 2018, 521-528.
- Tarigan, HD, Destiawati, F., & Fitriansyah, A. (2017). Implementation of C4.5 Algorithm Towards Customer Satisfaction. Ethos (Journal of Research and Community Service), Vol. 6, No, 80-86.
- Yunita, F. (2018). Application of Data Mining Using the K-Means Clustering Algorithm on Acceptance of New Students. Systemation, 7 (3), 238. <https://doi.org/10.32520/stmsi.v7i3.388>