INTERNATIONAL JOURNAL of DYNAMICS in ENGINEERING and SCIENCES (IJDES)
LLDIKTI WILAYAH X
Vol. 7, No. 2, Nov 2022
e-ISSN: 2502-0692

INCREASING CUSTOMER LOYALTY USING ALGORITHM METHODS K-MEANS

Elmi Rahmawati*[1], Vivi Puspita Sari[2,]
[1,2,3] Sistem Komputer, Universitas Putra Indonesia "YPTK" Padang, Indonesia

doi. 10.22216/jod.v7i1.1089
*Correspondence should be addressed to elmi_rahmawati@upiyptk.ac.id

Submitted :
15 May 2022

Accepted :
20 Sept 2022

Published :
1 Nov 2022

Abstract

Advances in technology allow humans to more quickly and precisely in carrying out various activities in everyday life. Along with the rapid progress technology, companies are required to continue to improve their ability to process more accurate information in order to meet their needs, desires, and levels of satisfaction so that customers remain loyal to using the company's products. The purpose of this research in general is to apply Web-based Customer Relationship Management with the K-means algorithm method on marketing/transactions at Nasution Family Doorsmeer. This research is a combined type of research, using descriptive and quantitative qualitative methods. From the research results, it is known that the existence of Web-based Customer Relationship Management can facilitate and assist in managing a good marketing strategy so as to increase sales revenue, and by providing the best service will encourage customer loyalty. and make it easier to determine loyal customers with the K-means algorithm. in order to continue to provide convenience to customers. 1. By implementing Customer Relationship Management with K-Means, you can build customer clustering to classify customers who use services so that service providers can identify the characteristics of their customers.

Keywords: Customer Relationship Management, K-means Algorithm, Loyalty, Web.

INTRODUCTION

Advances in technology allow humans to more quickly and precisely in carrying out various activities in everyday life. Along with the rapid progress technology, companies are required to continue to improve their ability to process more accurate information in order to meet their needs, desires, and levels of satisfaction so that customers remain loyal to using the company's products [1].

One of the companies engaged in the field of vehicle washing services is Doorsmeer Keluarga Nasution. Doorsmeer Family Nasution is located at Jl. Medan Padang Aek Godang Panyabungan, Mandailing Natal Regency. The Nasution Family Doorsmeer accepts motorized vehicle washing services such as car and motorcycle washing [2]. The Nasution Family Doorsmeer already has many customers and records it manually in a ledger book containing the name of the vehicle, the vehicle's license plate number, the date the washing was carried out, the officer who washed the vehicle and so on. Doorsmeer Family Nasution has implemented a strategy to build customer loyalty [3].

This strategy includes giving coupons to customers, and these coupons can be

exchanged after 10 washings. However, in determining this strategy, Doorsmeer has not based on a particular customer pattern. Here, we need a way to see loyal customers so that in determining strategies to build loyalty is more targeted. [4].

Customer Relationship Management (CRM) is a strategy that builds relationships to build customer loyalty or build long-term relationships to create greater value so as to be able to maintain market share and increase customer loyalty. CRM is a strategy on how to optimize profitability through developing customer satisfaction. This CRM strategy requires knowledge gained from customer data . One method that can be used in analyzing CRM is data mining. There are several data mining methods, namely prediction, association, decision tree, clustering and classification . With clustering, data patterns can be analyzed from naturally formed groups [5].

Data mining is a step in knowledge discovery in databases (KDD) which has a technique of analyzing data to explore hidden information in large and complex quantities, so as to produce output in the form of characteristics or patterns from the data. One of the data mining analysis techniques is cluster analysis, which is better known as clustering . Clustering is a data analysis method whose goal is to group data with the same characteristics into the same area [6]. One method of grouping data is non- hierarchical (border) which attempts to partition data into the form of two or more groups (clusters) with the same characteristics that are put into the same group [7].

The goal of Customer Relationship Management is to provide an optimal service and maintain existing relationships, know customer behavior, what are the needs needed by customers, because the key to a business is very dependent on how far we know about customers and what the needs are. required by customers to maintain good relations between customers and business owners [8][9][10].

RESEARCH METHODS

Preliminary Research

Preliminary research is the first step in conducting a study by first analyzing the problems to be developed. The purpose of problem identification is to find out about the problems that can be identified at the Nasution Family Doorsmeer, namely the process of giving promotions to loyal customers is still done manually, so it is very difficult for research objects to carry out this task because the promotion and discount must be given repeatedly. - Repeatedly due to inaccurate and precise calculation results or calculations that are not in accordance with the criteria and data owned by the customer, as a result the process takes quite a long time. Preliminary research is the first step in conducting research. By using the Customer Relationship Management approach with the K-Means algorithm method to provide promotions to customers who are entitled to receive them and provide solutions to problems that occur in the Nasution Family Doorsmeer.

Data collection

In this study, data was collected by interviewing the owners and their trustees who are engaged in the service sector. To get data about the Nasution Family Doorsmeer.

Analysis

Analysis is the process of finding, collecting and researching a problem from research conducted on the Nasution Family Doorsmeer.

Design

In this design stage, the author uses UML in designing the model for the system.

Implementation

In the implementation process, of course, it will involve the use of hardware (hardware) and software (software). The hardware (hardware) used .

Testing

This stage is the stage of testing the application that has been made, with the aim of knowing whether the application can be used or is able to run according to the design that has been done.

RESULTS AND DISCUSSION

Data analysis is the most important stage in the development of a system, data processing and assessment will find solutions in solving existing problems. This study uses data on the number of visits and total transactions in the Nasution Family Doorsmeer business, which can be seen from Table 1. below:

Name	Plate No	Number of Visits	Total Transactions
Aghna	BK 1264 QZ	3	210000
Adivas	BK 1263 ABL	4	240000
Alif	BK 1822 VN	2	120000
Lubis	BB 1446 FQ	7	420000
Dzaki	BK 2671NAHV	9	120000
Anisa	BM 9985 DI	1	70000
Novrida	BB 3280 RQ	10	138000
Grace	BK 1979 JBL	5	350,000
Nabila	BA 1168 QH	4	240,000
Aira	BK 1933 MX	1	50,000
Rara	BB 1752 AK	8	480,000
Ramli	BB 1079 EC	6	360,000
Sakdia	BK 888 MDL	10	700,000
Izza	BK 1752 AK	4	200,000
Jaza	BK 11 HSB	6	420,000
Flower	T 1425 DQ	10	600000
Holy	BB 1503 XR	2	140000
Budi	BK 4150 AGK	9	132000
Afiq	BA 1152 US	4	200000
Son	BK 1939 Y.D	2	120000

Table 1. Nasution Family Doorsmeer Transaction Data

Where the initial *centroid value* that has been determined is as follows:

1) *Clustering* Number of Transactions (X)
   a) *Centroids* 1 : 3
   b) *Centroids* 2 : 4
2) *Clustering* Total Transactions (Y)
   a) *Centroids* 1 : 210000
   b) *Centroids* 2 : 240000

After the initial *centroid value* has been determined, the next step is to enter the distance calculation stage for each data $d_i(x, y)$using formula 1 as follows:

$$d_i(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - C_x)^2 + (x_i - C_y)^2 : i = 1,2,3,\dots n} \quad (1)$$

Where d_i is the i-th data distance . x_i is the i-th data value . C_x the initial *centroid value* of Transaction Amount (C1,C2). C_Y is the initial *centroid value* of Total Transactions (C1, C2).

After all the data is calculated, allocate each data to the nearest *centroid* and determine the new *centroid value* using formula 2 as follows:

$$v = \frac{\sum_{i=1}^{n} x}{n} : i = 1,2,3 \dots \quad (2)$$

Where *v* is the new *centroid* . *n* is the number of data in the *cluster* .

After the *centroid value* is determined, the distance search process is repeated until the *cluster members* do not change from the previous iteration after being allocated.

First Iteration
The calculation of the object distance is carried out in the following way:

Centroids

$d_{1.1} = \sqrt{(3-3)^2 + (210000-210000)^2} = 0$
$d_{2.1} = \sqrt{(4-3)^2 + (240000-210000)^2} = 30000,00002$
$d_{3.1} = \sqrt{(2-3)^2 + (120000-210000)^2} = 90000,00001$
$d_{4.1} = \sqrt{(7-3)^2 + (420000-210000)^2} = 210000$
$d_{5.1} = \sqrt{(9-3)^2 + (120000-210000)^2} = 90000,0002$
$d_{6.1} = \sqrt{(1-3)^2 + (70000-210000)^2} = 140000$
$d_{7.1} = \sqrt{(10-3)^2 + (138000-210000)^2} = 72000,00034$
$d_{8.1} = \sqrt{(5-3)^2 + (350000-210000)^2} = 140000$
$d_{9.1} = \sqrt{(4-3)^2 + (240000-210000)^2} = 30000,00002$
$d_{10.1} = \sqrt{(1-3)^2 + (50000-210000)^2} = 160000$
$d_{11.1} = \sqrt{(8-3)^2 + (480000-210000)^2} = 270000$
$d_{12.1} = \sqrt{(6-3)^2 + (360000-210000)^2} = 150000$
$d_{13.1} = \sqrt{(10-3)^2 + (700000-210000)^2} = 490000,0001$
$d_{14.1} = \sqrt{(4-3)^2 + (200000-210000)^2} = 10000,00005$
$d_{15.1} = \sqrt{(6-3)^2 + (420000-210000)^2} = 210000$
$d_{16.1} = \sqrt{(10-3)^2 + (600000-210000)^2} = 390000,0001$
$d_{17.1} = \sqrt{(2-3)^2 + (140000-210000)^2} = 70000,00001$
$d_{18.1} = \sqrt{(9-3)^2 + (132000-210000)^2} = 78000,00023$
$d_{19.1} = \sqrt{(4-3)^2 + (200000-210000)^2} = 10000,00005$
$d_{20.1} = \sqrt{(2-3)^2 + (100000-210000)^2} = 90000,00001$

Centroid 2

$d_{1.2} = \sqrt{(3-4)^2 + (210000-240000)^2} = 30000,00002$
$d_{2.2} = \sqrt{(4-4)^2 + (240000-240000)^2} = 0$
$d_{3.1} = \sqrt{(2-4)^2 + (120000-240000)^2} = 120000$
$d_{4.2} = \sqrt{(7-4)^2 + (420000-240000)^2} = 180000$
$d_{5.2} = \sqrt{(9-4)^2 + (120000-240000)^2} = 120000,0001$
$d_{6.2} = \sqrt{(1-4)^2 + (70000-240000)^2} = 170000$
$d_{7.2} = \sqrt{(10-4)^2 + (138000-240000)^2} = 102000,00022$
$d_{8.2} = \sqrt{(5-4)^2 + (350000-240000)^2} = 110000$
$d_{9.2} = \sqrt{(4-4)^2 + (240000-240000)^2} = 0$
$d_{10.2} = \sqrt{(1-4)^2 + (50000-240000)^2} = 190000$
$d_{11.2} = \sqrt{(8-4)^2 + (480000-240000)^2} = 240000$
$d_{12.2} = \sqrt{(6-4)^2 + (360000-240000)^2} = 120000$
$d_{13.2} = \sqrt{(10-4)^2 + (700000-240000)^2} = 460000$
$d_{14.2} = \sqrt{(4-3)^2 + (200000-240000)^2} = 40000$
$d_{15.2} = \sqrt{(6-4)^2 + (420000-240000)^2} = 180000$
$d_{16.2} =$

$\sqrt{(10-4)^2+(600000-240000)^2}$ =360000,0001

$d_{17.2}$ = $\sqrt{(2-4)^2+(140000-240000)^2}$ =100000

$d_{18.2}$ = $\sqrt{(9-4)^2+(132000-240000)^2}$ =108000,0001

$d_{19.2}$ = $\sqrt{(4-4)^2+(200000-240000)^2}$ =40000

$d_{20.2}$ = $\sqrt{(2-4)^2+(100000-240000)^2}$ =120000

NO	C1	C2	CLUSTER
M1	0	30000,00002	C1
M2	30000,00002	0	C2
M3	90000,00001	120000	C1
M4	210000	180000	C2
M5	90000,0002	120000,0001	C1
M6	140000	170000	C1
M7	72000,00034	102000,0002	C1
M8	140000	110000	C2
M9	30000,00002	0	C2
M10	160000	190000	C1
M11	270000	240000	C2
M12	150000	120000	C2
M13	490000,0001	460000	C2
M14	10000,00005	40000	C1
M15	210000	180000	C2
M16	390000,0001	360000,0001	C2
M17	70000,00001	100000	C1
M18	78000,00023	108000,0001	C1
M19	10000,00005	40000	C1
M20	90000,00001	120000	C1

Table 2. First Iteration Object Distance

In table 2 all transaction data is directly allocated so that it can determine a new *centroid value* with the formula listed above by utilizing the ∑x value as the number of values in the *cluster members* and the n value as the number of members in *the cluster* . So the new *centroid value* in this iteration is as follows:

C1	4.272727273	136363,6364
C2	6,666666667	423333,3333

Table 3. First Iteration New *Centroid*

The next stage is to make a new iteration (second iteration) by repeating from the third stage in the previous iteration, it's just that the *centroid* used is the new *centroid* that was obtained in the previous *centroid* . After all stages are completed, the steps that need to be taken are to ensure the number of members in each *cluster is the* same as the number of members in each *cluster* in the previous iteration. When the state of the number of *clusters* is the same, then the process is stopped and *clusters are* formed. But if not, then the search will continue. The results of this study produced four iterations which can be seen as follows:

ITERATION	USED CENTROID VALUES				CLUSTER MEMBERS	
	C1		C2		1	2
	X	Y	X	Y		
1	3	210000	4	240000	11	9
2	4.272727273	136363,6364	6.666666667	423333,3333	13	7
3	4.230769231	152307,6923	7.428571429	475714,2857	13	7

Table 4. Clustering Results

CONCLUSION

Based on direct monitoring and analysis carried out on the Doorsmeer of the Nasution Family against the background of the descriptions and explanations described in the previous chapters, conclusions can be drawn including the following:

1. By implementing Customer Relationship Management with K-Means, you can build customer clustering to classify customers who use services so that service

providers can identify the characteristics of their customers .

2. With Customer Relationship Management with K-Means promotions can be directed to customers who are entitled to it.

3. With the implementation of Customer Relationship Management , reaching more new customers through a system that has been directly integrated with the internet network, so that new visitors or those who are not familiar with the Nasution Family Doorsmeer, both those who are around the business location and those who are far away, will find it easy to get to know the Nasution Family Doorsmeer through internet searches thereby increasing the number of customers who transact with the Nasution Family Doorsmeer business.

BIBLIOGRAPHY

[1] Zare, H., & Emadi, S. (2020). Determination of Customer Satisfaction using Improved K-means algorithm. Soft Computing, 24(22), 16947-16965.

[2] Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, *37*(7), 5259-5264.

[3] Nandapala, E. Y. L., & Jayasena, K. P. N. (2020, November). The practical approach in Customers segmentation by using the K-Means Algorithm. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (pp. 344-349). IEEE.

[4] Maryani, I., & Riana, D. (2017, August). Clustering and profiling of customers using RFM for customer relationship management recommendations. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-6). IEEE.

[5] Wu, J., Shi, L., Lin, W. P., Tsai, S. B., Li, Y., Yang, L., & Xu, G. (2020). An empirical study on customer segmentation by purchase behaviors using a RFM model and K-means algorithm. *Mathematical Problems in Engineering*, *2020*.

[6] Jenabi, G., & Mirroshandel, S. A. (2014). Using data mining techniques for improving customer relationship management. *European Online Journal of Natural and Social Sciences: Proceedings*, *2*(3 (s)), pp-3143.

[7] Wei, J. T., Lee, M. C., Chen, H. K., & Wu, H. H. (2013). Customer relationship management in the hairdressing industry: An application of data mining techniques. *Expert Systems with Applications*, *40*(18), 7513-7518.

[8] Rangriz, H., & Bayrami Shahrivar, Z. (2019). The impact of E-CRM on customer loyalty using data mining techniques. *BI Management Studies*, *7*(27), 175-205.

[9] Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, *36*(3), 4176-4184.

[10] Xiahou, X., & Harada, Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research*, *17*(2), 458-475.

[11] Anitha, P., & Patil, M. M. (2019). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University-Computer and Information Sciences*.