LOYALTY ASSESSMENT OF COMPANY COSTUMER WITH CLASSIFICATION METHOD

Fricles Ariwisanto Sianturi¹, Jonson Manurung² R.Fanry Siahaan³

Informatics Management Study Program¹ Software Engineering Study Program² Technical Informaric Study Program³

^{1,2,3}STMIK Pelita Nusantara Medan, North Sumatra, Indonesia sianturifricles@gmail.com¹, jhonson.geo@gmail.com², rfanry@gmail.com³

Abstract

Companies in general want the customers they have to be able to sustain forever. To make this happen is not something that is easy in the current climate of intense business competition, considering that there are rapid changes that can occur at any time, such as changes in customers, competitors and changes in broad conditions that are always dynamic. This requires policy makers to develop a strategy capable of achieving sales growth targets, increasing the company's market share. For this reason, an analysis of customer loyalty is needed. For this reason, an analysis is needed to understand and assess customer loyalty using a classification design method. With classification, information can be produced more quickly and the information presented is analytical in nature so that it is easy to use for decision making.

Keywords: Data Mining, Loyalty, classification

1. Introduction

PT. Telkom's Medan branch is a state-owned company engaged in telecommunications services with a variety of products they have. The products owned by PT. Telkom, among others, Telkom Flexi for home, Telkom Home Phone for landlines with cable, as well as internet service packages with the name Speedy. PT. Telkom, especially the Medan branch, already has many consumers (customers), individuals and companies that use telecommunications services owned by PT. Telkom, such as the Speedy Package and telephone by cable. But currently PT. Medan Branch Telkom has a big problem, namely the emergence of service providers and services that have the same line of business as those owned by PT. Telkom, as mobile phone marketing is increasing, there are also internet service providers, up to cellular providers such as Telkomsel, XL and Indosat.

PT. Telkom must be able to make decisions about what to take as a result of the emergence of service and service providers that have the same line of business which are increasingly widespread and increasingly out of control. This in turn will affect customer loyalty to PT. Telkom. This can be seen from the decline in customers who use the services of PT. Telkom, as an example of using cable telephones (landlines). Since the emergence of other communication products such as cellular telephones, many customers of PT. Telkom which switched to cellular telephone products which are considered simpler and can be carried anywhere. Another example is that many companies have moved to other internet service companies. This example is a condition of decreasing loyalty and trust of some people towards PT. Telkom. On the other hand, there are still many parties who are still loyal to use products from PT. Telkom.

Classification is the process of finding a model or function that describes or differentiates a concept or data class, with the aim of being able to estimate the class of an object whose label is unknown. This model can take the form of a decision tree[1]





2. Literature Review

2.1 Customer loyalty

Customer loyalty is customer loyalty to a product and service. With customer loyalty, it can guarantee the survival (continuity) of the company in the long run. Basically, loyalty is a customer's decision to voluntarily, continuously subscribe to a certain company for a long period of time. Loyalty will continue as long as the customer feels and is able to receive better value (including higher quality in terms of price) than would be obtained by switching to another supplier.[2]

2.2 Data Mining

Data mining or often referred to as knowledge discovery in database (KDD) is an activity that includes the collection, use of historical data to find regularities, patterns or relationships in large data. This data mining expenditure can be used to help make decisions in the future. The development of KDD causes the use of pattern recognition to decrease because it has become part of data mining[3]

The terms data mining and knowledge discovery in database (KDD) is often used take turns to explain the excavation process hidden information in a database big. In fact, both terms have a different concept, but it has interrelation with each other, which is where the stages throughout the Knowledge Discovery in Database (KDD) is data mining

KDD process sizes are as follows:[3]

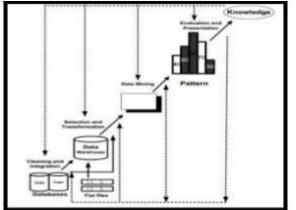


Figure 1. Knowledge Discovery stage on Data Mining (KDD)

2.3 Classification

Classification is the process of finding a model or function that describes or differentiates a concept or data class, with the aim of being able to estimate the class of an object whose label is unknown. This model can take the form of a decision tree.[4] [5]

The classification process is usually divided into two phases: learning and testing. In the learning phase, some of the data that the data class has known is fed to form a prediction model.[5]Because it uses data that has been pre-labeled by an expert in the field as examples of correct data, classification is often referred to as a supervised method. Then in the test phase the model that has been formed is tested with some other data to determine the accuracy of the model. If the accuracy is sufficient, this model can be used to predict unknown data classes[6].

The stages of classification in data mining consist of:

- 1. Development of the model, in this stage a model is created to solve the class or attribute classification problem in the data. This model is built based on the training set an example of data from the problems faced, this training set already has complete information both in terms of attributes and classes.
- 2. Application of the model, at this stage the previously built model is used to determine the attributes / class of a new data whose attributes / class are not previously known Evaluation, at this stage the





results of applying the model in the previous stage are evaluated using measured parameters to determine whether the model is acceptable.[7]

In designing a classification method, the process is changing the form of data into a tree model, changing the tree model into a rule, and simplifying the rules. In general, the Classification method for building a decision tree is as follows:[8]

1. Select the attribute as the root

2. Create a branch for each value

3. Divide cases into branches

4. Repeat the process for each branch until all cases on the branch have the same class

To choose an attribute as the root, it is based on the highest gain value of the existing attributes. To calculate gain, a formula is used as shown in equation 1 below:[9]

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

Information:

S: case setA: attributen: the number of partitions attribute A|Si|: the number of cases on the ith partition|S|: number of cases in S

Meanwhile, the calculation of the entropy value can be seen in equation 2 below:

$$Entropy(S) = \sum_{i=1}^{n} - pi * \log 2 \ pi \ (2)$$

Information:

S: case set

A: features

n: number of partitions S

Pi: the proportion of Si to S

To determine the entropy and gain values, calculations are performed manually to get the root and branch nodes based on the entropy and gain values.[10][11]

3. Results and Discussion

The Classification method in the process requires criteria that will be used as calculation material in the ranking process. The criteria that are taken into consideration in the customer loyalty assessment process are as shown in table 1.

Table 1. Criteria

Criteria	Information
C1	Type of Customer
C2	Type of Customer Using Telkom Products
C3	Length of Subscription
C4	Big Dues
C5	Many Products

Customer Type Entropy Value

70 70 10 10 Entropy (Long) = (- - * Log 2 (-)) + (- - * Log 2 (-))

$$\bigcirc \bigcirc \bigcirc \bigcirc$$

 $\begin{array}{cccc} 80 & 80 & 80 & 80 \\ \text{Entropy (Long)} &= (0.23) + (0.004) \\ \text{Entropy (Long)} &= 0.23 \\ & 20 & 20 & 0 & 0 \\ \text{Entropy (New)} &= (---) * \text{Log 2}(--)) + (--- * \text{Log 2}(--)) \\ & 20 & 20 & 20 & 20 \\ \text{Entropy (New)} &= 0 \end{array}$

Entropy Value Using Telkom Products

Entropy (Yes) = (- * Log 2) + (- * Log 2 (-))Entropy (Yes) = (0.23) + (0.003)Entropy Yes Yes = 0.2324.2 Entropy (No) = (- * Log 2 (-)) + (- * Log 2 (-))Entropy (No) = (0.24) + (0.001)Entropy (No) = 0.24

Payment Type Entropy Value

Entropy (Already paid) = (- * Log 2 (-)) + (- * Log 2 (-))Entropy (Already paid) = 0Entropy (Already paid) = (- * Log 2 (-)) + (- - * Log 2 (-))Entropy (Already paid) = (0.17) + (0.018)Entropy (Already paid) = (0.19)

Entropy Value of Contribution

Entropy (Business) = (- * Log 2 (-)) + (- * Log 2 (-))Entropy (Business) = (0.28) + (0.0001)Entropy (Business) = (0.28)Entropy (Resident) = (- * Log 2 (-)) + (- - * Log 2 (-))Entropy (Resident) = (0.19) + (0.009)Entropy (Resident) = 0.2Entropy (Pem_TNI_Polri) = (- * Log 2 (-)) + (- * Log 2 (-))Entropy (Pem TNI Polri) = 0The Value of the Multi-Product Entropy Entropy (Consumptive) = (- - * Log 2 (-)) + (- - * Log 2 (-))Consumptive Entropy) = 0



34 34 0 0 Entropy (less consumptive) = (- * Log 2 (-)) + (- * Log 2 (-))34 34 34 34 Entropy (Less Consumptive) = (0) 8 10 8 10 Entropy (Not Consumptive) = (- * Log 2 (-)) + (- * Log 2 (-))18 18 18 18 Entropy (not consumptive) = (0.06) + (0.1)Entropy (not consumptive) = 0.16As for the Gain value for each attribute line, it is calculated using equation 1, as shown below: 80 20 Gain (Customer Type) = 0.24 - ((- * 0.23) + (- - * 0)) = 0.06100 100 Gain (Using Telkom Products) = 0.24 - ((- * 0.23) + (- * 0)) = 0.13100 100 40 60 Gain (Type of Payment) = 0.24 - ((- * 0) + (- * 0.19)) = 0.16100 100 Gain (Output Magnitude) 49 6 = 0.24 - ((- * 0.28) + (- * 0.2) + (- * 0.2) + (- * 0))100 100 100 = 0.19 $48 \quad 34 \quad 18$ Gain (Number of Products) = 0.24 - ((--*0) + (--*0) + (--*0) = 0.21 100 100 100

1. Calculating the number of cases, the number of cases for Loyal decisions, the number of cases for Disloyal decisions, and Entropy from all cases and cases divided by the attributes of customer types, Using Telkom Products, types of payment, Contribution, and Number of Products. After that, do the gain calculation for each attribute. The calculation results are shown in table 2.

Table 2. Node 1 calculationNumber							
			of cases	Loyal	Disloyal	F (<i>a</i> .
No.		1	(S)	(S1)	(S2)	Entropy	Gain
1	TOTAL		100	90	10	0.24	
	CUSTOMER						
	TYPES						0.06
		Long	80	70	10	0.23	
		New	20	21	0	0	
	USING						
	TELKOM'S						
	PRODUCTS						0.13
		Yes	74	66	8	0.23	
		Not	26	24	2	0.24	

JURNAL INFOKUM, Volume 9, No.1 Desember 2020 ISSN : 2302-9						
TYPE OF						
PAYMENT						0.16
	Already paid	60	60	0	0	
	Not yet paid	40	30	10	0.19	
GREAT DUE						0.19
	Business	45	44	1	0.28	
	Resident	59	40	9	0.2	
	Pem_TNI_Polri	6	6	0	0	
MANY						
PRODUCTS						0.21
	Consumptive Less	48	48	0	0	
	Consumptive Not	34	34	0	0	
	Consumptive	18	8	10	0.16	

http://infor.seaninstitute.org/index.php/infokum/index

Entropy Value Using Telkom Products Entropy Using Telkom Products

Entropy (Yes) = $\begin{pmatrix} 7 & 7 & 10 & 10 \\ - & Log 2 & (-) + (- & Log 2 & (-)) \\ 17 & 17 & 17 & 17 \end{pmatrix}$ 10 Entropy (Yes) = (0.05) + (0.010)Entropy Yes = 0.15Entropy (No) = (- - * Log 2 (-)) + (- * Log 2 (-))1 1 1 1 1 1 1Entropy (No) = 0

Customer Type Entropy Value

11 10 10 Entropy (Old) = (- * Log 2 (-)) + (- * Log 2 (-))11 11 11 11 Entropy (Old) = (0.02) + (0.024)Entropy (Old) = 0.2477 0 0 Entropy (New) = (- - * Log 2 (-)) + (- - * Log 2 (-))7 7 7 7 7 Entropy (New) = 0

Payment Type Entropy Value

1 11 10 10 Entropy (Unpaid) = (- - * Log 2 (-)) + (- - * Log 2 (-))11 11 11 11 Entropy (Unpaid) = (0.002) + (0.24)Entropy (Unpaid) = 0.2477 0 Entropy (Already Paid) = (- * Log 2 (-)) + (- * Log 2 (-))7 7 7 7



Entropy (already paid) = 0

Entropy Value of Contribution

2. Calculating the number of cases, the number of cases for loyal decisions, and the number of cases for disloyal decisions, and the entropy of all cases and cases divided by the customer type attribute which can be the root node of the non-consumptive attribute value. After that, do the Gain calculation for each attribute. The calculation results are shown in table 2.

No.			Number of Cases (S)	Loyal (S1)	Disloyal (S2)	Entropy	Gain
1.1	MANY PRODUCTS (not consumptive) CUSTOMER TYPES		18	8	10	0.16	0.02
		Long New	11 7	1 7	10 0	0.24 0	0.02
	USING TELKOM'S PRODUCTS						0.02





http://infor.seaninstitute.org/index.php/infokum/index
JURNAL INFOKUM, Volume 9, No.1 Desember 2020

ISSN: 2302-9706

	Yes	17	7	10	0.15	
	Not	1	1	0	0	
TYPE OF						
PAYMENT						0.02
	Already paid	7	7	0	0	
	Not yet paid	11	1	10	0.24	
GREAT DUE						0.02
	Business	0			0	
	Resident	18	8	10	0.14	
	Pem_TNI_Polri	0	0	0	0	

From the results in table 2, it can be seen that each attribute has the same gain which is worth 0.02. So to make the root node of the many products (not consumptive), it is seen from the attributes that most influence customer loyalty, namely the type of customer. Thus the type of customer can be used as the root node, there are 2 attribute values of the customer type, namely old and new. The new attribute value has classified cases with the highest value of 0, namely the risk decision is loval, so there is no need for further calculations. However, the old attribute values still need to be calculated again.

Entropy Value Using Telkom Products

10	10	0	0
Entropy (Yes) = $(- * \text{Log } 2)$	2 () + (- * Log	2 ())
10	10	10	10
Entropy (Yes) $= 0$			
1 1	0	0	
Entropy (No) = $(* \text{Log } 2)$		* Log	2 ())
1 1	1	1	
Entropy (No) $= 0$			

Payment Type Entropy Value

Entropy (Unpaid) = (- * Log 2(-)) + (- * Log 2(-)) = 00 Entropy (Unpaid) = 0 Entropy (Unpaid) = (- * Log 2(-)) + (- * Log 2(-))11 11 11 11 10 Entropy (Unpaid) = (0.002) + (0.24)Entropy (Unpaid) = 0.24

Entropy Value of Contribution

 $\begin{array}{c} 0 & 0 & 0 & 0 \\ \text{Entropy (Business)} = (- - * \text{Log 2} (-)) + (- - * \text{Log 2} (-)) \\ 0 & 0 & 0 \end{array}$ Entropy (Business) = 0Entropy (Resident) = (0.02) + (0.024)Entropy (Resident) = 0.24Entropy (Pem_TNI_Polri) = $(-- * \log_{0} 2 (-)) + (- * \log_{0} 2 (-))$ Entropy (Pem_TNI_Polri) = 0



As for the Gain value for each attribute line, it is calculated using equation 1, as shown below:

Gain (Using Telkom Products) 34 10 = 0.24 - ((- * 0) + (- * 0))11 11 Gain (Using Telkom Products) = 011 0 Gain (Subscription Period) = 0.24 - ((- * 0) + (- - * 0))11 11 Gain (Subscription Period) = 0 $\begin{array}{ccc} 11 & 0 & 0\\ \text{Gain (Number of Products)} = 0.24 - ((- - * 0) + (- - * 0) + (- - * 0) \end{array}$ 11 11 11 Gain (Multiple Products) = 0

3. Calculates the number of cases, the number of cases for loyal decisions, and the number of cases for disloyal decisions, and the Entropy of all cases and cases divided by the Multiple Product (non-consumptive) attribute, the type of customer that can be a branch node of the old attribute value. After that, do the Gain calculation for each attribute. The calculation results are shown in table 4.

			number				
			of cases	Loyal	Disloyal	_	
No.			(S)	(SI)	(S2)	Entropy	Gain
	MANY						
	PRODUCTS						
	(not consumptive)						
	CUSTOMER						
1.0	TYPES				10	0.04	
1.2	(long)		11	1	10	0.24	
	USING TELKOM'S						0
	PRODUCTS	Yes	10	0	10	0	0
						0	
		Not	1	1	0	0	
	TYPE OF						
	PAYMENT						0
		Already					0
		paid	0	0	0	0	
		Not yet	Ŭ	Ū	0	0	
		paid	11	1	10	0.24	
		P		-			
							0
	GREAT DUE	Dereineren	0	0	0	0	0
		Business	0	0	0	0	
		Resident	11	1	10	0.24	
		TNI_	0	0	0	0	

Table 4. Calculation of Node 3





Police

From the results of table 4. the highest gain is obtained for the attribute Using Telkom Products, the type of payment, the Contribution of 0. So that the calculation is stopped.

4. Conclusion

From the experimental results that the author conducted on this research, the writer can draw some conclusions related to the research process and the content of the research itself.

- 1. The process must be carried out with complete data so that the assessment will be carried out objectively.
- 2. The application of classification methods can facilitate and provide settlement calculations and criteria in assessing customer loyalty.

Reference

- E. Elisa, "Analysis and Application of the C4.5 Algorithm in Data Mining to Identify Factors [1] Causing PT.Arupadhatu Adisesanti Construction Accidents," J. Online Inform., vol. 2, no. 1, p. 36, 2017, doi: 10.15575 / join.v2i1.71.
- S. Normasari, S. Kumadji, and A. Kusumawati, "The Influence of Service Quality on Customer [2] Satisfaction, Company Image and Customer Loyalty," J. Adm. Business, 2013.
- M. Fahmi and FA Sianturi, "ANALYSIS OF APRIORIC ALGORITHM ON CONSUMER [3] ORDERING AT CAFÉ THE L. COFFE COFFEE, "SAINTEK (Journal of Science and Technology., vol. 1, no. 1, pp. 52–57, 2019.
- A. Mukminin and D. Riana, "Comparison of C4 Algorithms. 5, Naïve Bayes And Neural [4] Network For Soil Classification, "J. Inform., 2017.
- FA Sianturi, "Decision Tree Analysis in Student Data Processing," MEANS (Media Inf. Anal. [5] And Sist. vol. 3. no. 2, pp. 166–172, 2018. [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/.
- J. Simarmata et al., "Multimedia of number recognition for early childhood using image object," [6] Int. J. Eng. Technol., Vol. 7, no. 3.2 Special Issue 2, pp. 796-798, 2018, doi: 10.14419 / ijet.v7i3.2.18760.
- YI Kurniawan, "Comparison of Naive Bayes Algorithm and C.45 in Data Classification Mining, [7] "J. Technol. Inf. and Computer Science., 2018, doi: 10.25126 / jtiik.201854803.
- FA Sianturi, B. Sinaga, PM Hasugian, T. Informatics, and S. Utara, "Fuzzy Multiple Attribute [8] Decisison Macking Using Oreste Method to Determine Promotion Location," J. Inform. Pelita Nusant. vol. 3, no. 1, pp. 63–68, 2018, [Online]. Available: http://ejurnal.pelitanusantara.ac.id/index.php/JIPN/article/view/289.
- [9] DS Kusumo, MA Bijaksana, and D. Darmantoro, "DATA MINING WITH APRIORIC ALGORITHM IN ORACLE RDBMS," TECHNOLOGY - J. Researcher. and Pengemb. Telecomun. Control, Computer, Electr. and Electrons., 2016, doi: 10.25124 / tektrika.v8i1.215.
- EP Cynthia and E. Ismanto, "The C.45 Algorithm Decision Tree Method in Classifying Data on [10] Sales of Fast Food Store Business," Jurasik (Jurnal Ris. Sist. Inf. And Tek. Inform., vol. 3, no. July, p. 1, 2018, doi: 10.30645 / jurasik.v3i0.60.
- [11] IH Witten, E. Frank, MA Hall, and CJ Pal, Data Mining: Practical Machine Learning Tools and Techniques. 2016.



30