

# Customer Satisfaction Analysis Using the C4.5 Algorithm at Askara Minimarket

Toto Hermanto

Universitas Utpadaka Swastika, Tangerang, Indonesia, 15112  
E-mail: totohermanto8@gmail.com

Accepted: April 28, 2024 | Published: April 30, 2024

## ABSTRACT

Customer satisfaction is one way to improve the quality of service in a company, service satisfaction is also a factor in the company's development. If the company does not have quality service to the company, the company will not be able to develop and compete with other companies. The relationship between the company and consumers must run well, before employees enter the shop area, employees must be trained first. After the consumer has finished shopping, the consumer must choose whether he is satisfied or dissatisfied in the application, the results of the consumer's satisfaction and dissatisfaction will be entered into the data, this data will later be processed into Data Mining. The C4.5 algorithm is a data classification algorithm using a decision tree technique that is well known and preferred because it has advantages. This research process uses Weka software to create a decision tree from the test results and it is hoped that the company can improve its services to customer satisfaction. From the test results using Weka software and using data obtained from questionnaires with customer satisfaction from 50 respondents, the results obtained were 0.0963, which indicates that consumers are satisfied with the minimarket service and from the results of the Weka test after visualizing the tree, the service is good but the place is lacking.

**KEYWORDS:** Customers and companies; Data Mining; Algorithm C4.5

## 1. Introduction

Every year technological developments continue to develop, with increasingly developed technology, the increasingly advanced development of the retail world causes companies to have to face intense competition. Companies in general want customers they can have forever. To make repeat customers loyal customers is not easy, therefore companies are increasingly improving quality and productivity [1]. Human Resources must be able maximized well. Knowledge - the knowledge possessed by human resources is a source of company development. Companies must treat employees and the company well, therefore employees must treat customers like kings, treat customers in a friendly and polite manner so that customers can shop again. The company's awareness of the importance of

assessing the company is the most important factor so that the company can survive as long as possible. for a long time and opened branches everywhere.

Maintaining loyal customers is not easy, companies have to do a lot, for example employees are required to innovate to make the company better. Consumers often experience dissatisfaction with the company, if the service is less than optimal, there is no greeting to the consumer then the consumer feels uncomfortable and feels less than optimal in service. However, if consumers feel well served and are greeted with greetings, then consumers themselves will feel satisfied. Therefore, consumers are kings who must be served as well as possible.

According to customer satisfaction surveys, consumers who fill out customer satisfaction forms. The data will be analyzed

using data mining using the C4.5 algorithm because this algorithm can be understood by users who want to research the results of customer satisfaction. The C4.5 algorithm is a data classification algorithm using a well-known and understandable decision tree technique. These advantages include, for example, being able to manage numeric and discrete data, being able to handle missing attribute values, which produces rules that are easy to interpret and the fastest among other algorithms. Prediction accuracy is the ability of the model to be able to predict class labels for new data or data that was not previously known well.

## 2. Theoretical Foundation

### 2.1 Service Quality

Service is something that is very important for consumers, to create consumer satisfaction. If the company provides the best service to consumers, consumers will not be discouraged. How good the level of service provided to consumers is one of the success factors in competition [1].

### 2.2 Consumer Satisfaction

Consumer satisfaction is an important thing in a company, where consumers are the spearhead of a company to achieve success, services that can satisfy consumers will provide and increase the company's competitiveness.

### 2.3 Data Mining

Data mining provides the results of research on the problems being faced by the company then produces large data and then creates certain rules, patterns or models to recognize new data that is not in the stored data rows. Data mining carries out a search

process by looking for information or sources that will be used as research [2]

### 2.4 Data mining stages

As a series of data mining processes it can be divided into several stages.

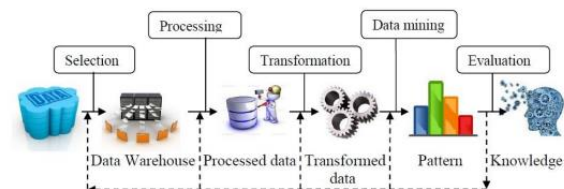


Figure 1. Data mining stages

#### A. Data cleaning

Data cleaning is the process of cleaning inconsistent data. Data is taken based on a database or from information.

#### B. Data Selection

Not all of the data taken from applications and data can be used, the information obtained to produce solutions is taken from the database and then the data is analyzed so that the data can be useful for the company.

#### C. Data integration

Integration in data is combining data from an old database into a new database. Often we hear about databases, databases are data taken from applications.

### 2.5 Decision Trees

A decision tree is called a decision tree, describing a sequence of tree-like structures where each node represents a class. Decision trees are called roots, decision trees are one of the classification methods most often used for classification and prediction. This is supported by the ability to describe the structure relatively quickly, the results of the model formed are very easy to understand.

## 2.6 Algorithm C4.5

The C4.5 algorithm is an algorithm for converting large facts into a decision tree that represents rules. Some of the developments made in C4.5 include being able to overcome missing values, continuous data, and pruning [3]. The formula used in calculations with the algorithm can be explained as follows:

$$\text{Gains}(S, A) = \text{Entropi}(S) - \sum_{i=1}^n \left( \frac{|S_i|}{|S|} \right) \times \text{Entropi}(S_i) \dots\dots\dots (1)$$

Figure 2. C4.5 algorithm calculation

Where :

A: Attributes

S: Case Set

n: Number of partitions of attribute A

|S<sub>i</sub>| : Number of cases in the i<sup>th</sup> partition

|S| : Number of cases in S

## 3. Design Concept

### 3.1 Data collection process

Data was collected using a questionnaire from 50 respondents, and used in January 2024 to fill out the questionnaire. Questionnaires filled out by consumers, regarding service and satisfaction. Each attribute has a value for each attribute to determine satisfaction and dissatisfaction. An answer from consumers is calculated based on the resulting value [4].

NO	NAMA	PELAYANAN	KEBERSIHAN	TEMPAT	HARGA
1	ANDI	BAIK	PEDULI	STRATEGIS	MURAH
2	LAUREN	BAIK	PEDULI	STRATEGIS	MURAH
3	TONI	BAIK	PEDULI	MACET	MAHAL
4	UKI	BAIK	KURANG PEDULI	KURANG STRATEGIS	STANDAR
5	SALSA	CUKUP	PEDULI	STRATEGIS	MURAH
6	DIMAS	BAIK	KURANG PEDULI	KURANG STRATEGIS	MURAH
7	AMI	CUKUP	PEDULI	STRATEGIS	MURAH
8	SUSAN	BAIK	PEDULI	STRATEGIS	STANDAR
9	ANGGA	BAIK	PEDULI	MACET	STANDAR
10	YOGI	BAIK	KURANG PEDULI	KURANG STRATEGIS	MAHAL
11	SUSI	KURANG BAIK	PEDULI	STRATEGIS	STANDAR
12	KIKI	BAIK	PEDULI	KURANG STRATEGIS	MURAH
13	DONI	KURANG BAIK	KURANG PEDULI	STRATEGIS	MURAH
14	ANGGI	CUKUP	PEDULI	STRATEGIS	MURAH
15	UMI	BAIK	PEDULI	KURANG STRATEGIS	STANDAR
16	FAQI	BAIK	KURANG PEDULI	STRATEGIS	STANDAR
17	DODO	BAIK	PEDULI	KURANG STRATEGIS	MAHAL
18	UMAY	CUKUP	KURANG PEDULI	STRATEGIS	STANDAR
19	UMU	KURANG BAIK	PEDULI	KURANG STRATEGIS	MURAH
20	NITA	CUKUP	PEDULI	MACET	MURAH
21	INTAN	BAIK	PEDULI	STRATEGIS	MURAH
22	RIA	CUKUP	PEDULI	KURANG STRATEGIS	MURAH

23	SITI	BAIK	PEDULI	KURANG STRATEGIS	MURAH
24	ANTO	KURANG BAIK	PEDULI	STRATEGIS	STANDAR
25	DELLA	BAIK	PEDULI	KURANG STRATEGIS	STANDAR
26	DIAN	KURANG BAIK	PEDULI	STRATEGIS	STANDAR
27	ENI	BAIK	PEDULI	STRATEGIS	STANDAR
28	YUNI	BAIK	PEDULI	KURANG STRATEGIS	STANDAR
29	SUSAN	BAIK	PEDULI	MACET	MURAH
30	PAMELA	KURANG BAIK	PEDULI	STRATEGIS	MURAH
31	YANI	CUKUP	PEDULI	KURANG STRATEGIS	MURAH
32	APRILIA	BAIK	PEDULI	STRATEGIS	STANDAR
33	TOMI	BAIK	KURANG PEDULI	MACET	MAHAL
34	TOMAS	KURANG BAIK	PEDULI	STRATEGIS	STANDAR
35	DIDI	BAIK	PEDULI	KURANG STRATEGIS	MURAH
36	DONI	CUKUP	KURANG PEDULI	STRATEGIS	MURAH
37	PINKAN	BAIK	PEDULI	STRATEGIS	MURAH
38	ELISA	BAIK	KURANG PEDULI	KURANG STRATEGIS	MURAH
39	AKMAL	BAIK	PEDULI	MACET	MURAH
40	RUSLI	CUKUP	KURANG PEDULI	KURANG STRATEGIS	STANDAR
41	DAMAR	CUKUP	PEDULI	STRATEGIS	STANDAR
42	PUTRA	BAIK	PEDULI	MACET	MAHAL
43	IMAS	CUKUP	PEDULI	KURANG STRATEGIS	STANDAR
44	ANGGORO	BAIK	KURANG PEDULI	STRATEGIS	MAHAL
45	PUTRI	KURANG BAIK	KURANG PEDULI	KURANG STRATEGIS	STANDAR
46	JOJO	CUKUP	PEDULI	STRATEGIS	STANDAR
47	TONO	BAIK	KURANG PEDULI	KURANG STRATEGIS	MURAH
48	TUTI	KURANG BAIK	PEDULI	STRATEGIS	MAHAL
49	JOHAN	BAIK	KURANG PEDULI	MACET	MURAH
50	MELISA	CUKUP	PEDULI	KURANG STRATEGIS	MAHAL

Figure 3: Questionnaire Result Data

The data from the questionnaire results will later be used as research to measure the level of consumer satisfaction using the C4.5 algorithm [5]. After analyzing the data, 50 respondents filled out the questionnaire.

### 3.2 Test Results

Test results using weka software.

```

Classifier output
PELAYANAN = BAIK: KURANG STRATEGIS (28.0/16.0)
PELAYANAN = CUKUP
| NO <= 21: STRATEGIS (5.0/1.0)
| NO > 21: KURANG STRATEGIS (8.0/3.0)
PELAYANAN = KURANG BAIK: STRATEGIS (9.0/2.0)

Number of Leaves : 4

Size of the tree : 6

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      17      34 %
Incorrectly Classified Instances    33      66 %
Kappa statistic                    -0.0963
Mean absolute error                 0.4261
Root mean squared error             0.5275
Relative absolute error             102.5407 %
Root relative squared error         115.7026 %
Total Number of Instances          50

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.522  0.491  0.480  0.522  0.500  -0.040  0.523  0.526  STRATEGIS
0.000  0.119  0.000  0.000  0.000  -0.145  0.494  0.169  MACET
0.263  0.484  0.250  0.263  0.256  -0.219  0.377  0.315  KURANG STR
Weighted Avg.  0.340  0.424  0.316  0.340  0.327  -0.068  0.461  0.389
    
```

Figure 3. Test Results

From the test results using Weka software and using data obtained from questionnaires with customer satisfaction from 50 respondents, the results obtained were 0.0963, which indicates that consumers are satisfied with minimarket services.

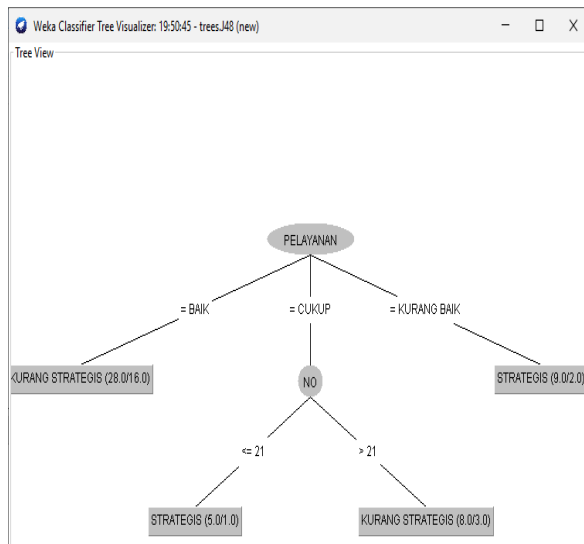


Figure 4: Visualize tree

From the results of the Weka test after the visualize tree results, it is explained that:

- Good service but less strategic location (28.0/16.0)
- Service is not good but strategic location (9.0/2.0)
- Service is adequate but strategic (5.0/1.0)
- Adequate service but less strategic (8.0/3.0)

#### 4. Conclusion

Based on the results, it can be concluded that the C4.5 Algorithm can be used as a classification method that can analyze customer satisfaction in serving consumers. The decision tree formed produces 2 rules which become a reference for customers to consumers. Even though the location is less strategic but service is used as a reference for service quality, so consumers are very happy to shop at Askara minimarkets.

#### REFERENCES

- Naya Chandra, Siswandi Arif, "CUSTOMER SATISFACTION ANALYSIS USING THE C4.5 ALGORITHM AT PT. CAHAYA INDOTAMA ENGINEERING," vol. 13, pp 2407 - 3903, 2023
- Azwati Nurul, Elisa Erlin, "Consumer Satisfaction Analysis Using the C4.5 Algorithm," pp 978-602-52829, 2020
- Adriana Mezi, Yulianti Liza, Elfianty Lena, "Customer Satisfaction Analysis Using the C4.5 Algorithm" Vol. 7, pp 2548-1916, 2022
- Kurniawati Ika, Telaumbanua Destalmawati, "Application of the C4.5 Algorithm for classification of customer satisfaction in delivery services" Vol.06, pp.001-006
- Aditya Nugroho1 Caesar Rizky, Kristiana Titin, "Application of the C4.5 Algorithm for Customer Satisfaction at the Chantik Perfume Online Store" Vol. 3, pp 10-21, 2022
- Putri Maisa Sherly, Arnomo Ani Sasa, "Application of the C4.5 Algorithm to Predict Service Quality on Consumer Satisfaction (Case Study: Hinet Batam)" Vol. 1, pp 2686 – 228, 2020
- Rezza Fauzy, Winanjaya Riki, Susiani "Analysis of customer satisfaction levels by applying the C4.5 algorithm" vol.2 pp 2774-3659, 2022
- Purwaningsih Esti, Nurelasari Ela "Application of the C4.5 algorithm in measuring the level of customer satisfaction with fast food restaurant services" vol.3, pp 2747-1779, 2022
- Edsel Barito, Beng Tji Jap, Arisandi Desi "Application of the C4.5 Algorithm for the classification of students receiving Covid-19 social assistance" vol.2, pp 2876-1886, 2022
- Azizah Nur Adilla, Purnamasari Irma Ade, Ali Irfan "Application of the C4.5 algorithm to predict the stock of beverage ingredients in café semanis" vol.1, 2024