Jurnal Mantik
Volume 4, Number 2, Aug 2020, pp. 1076-1084 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik

Implementation Of C4.5 Algorithm To Analyze Library Satisfaction Visitors

Fristi Riandari[1], Hengki Tamando Sihotang[2]

STMIK Pelita Nusantara Medan Jln. Iskandar Muda No. 1, 20154, Indonesia

Email: fristy.rianda@ymail.com

ARTICLE INFO

Article history:
Received: 03-07-2020
Revised: 09-07-2020
Accepted: 30-07-2020

Keywords:
Visitor Satisfaction, Library, Data Mining, C4.5

ABSTRACT

Measuring visitor satisfaction, especially in the library, is very important to note considering the library is a means intended to help the academic process or add insight to visitors. Measuring library visitor satisfaction is very important to note whether the services expected by visitors are in accordance with what is received, which will greatly assist the university in processing the library so that it can be used properly, so that the books provided can be of maximum use and not only the quality of service provided for the library provided can be an attraction for visitors to become fond of visiting the library reading books that will help academically the average visitor is a student. The system needed is the Data Mining Application in Measuring Satisfaction of library visitors at STMIK Pelita Nusantara. Where library visitors will be objects that provide an assessment / opinion on variables that have characteristics, namely Tangiable, Reability, Responsivnes, Assurance, Empathy. This system was built with the C4.5 algorithm and the Testing system with the application RapidMiner Studio 9.7.

1. Introduction

C4.5 Algorithm is a case solving solution that is often used to make decision trees in problem solving in classification techniques that have characteristics, namely the process of determining entropy and gain values. A study explains that the C4.5 algorithm is a classification technique using entropy and information advantage as a separator in a decision tree (Florence A.M and Savithri. R., 2013). A research comparing the C4.5 algorithm and CART algorithm in classifying student scores explained that the C4.5 algorithm has a higher accuracy value of 85.61% while the CART algorithm has an accuracy value of 84.95% (Rahmayuni I., 2014). In other studies, data mining was implemented to measure the length of a student's study and the results of testing were obtained which explained that the error rate in measuring the length of a student's study was only 5% (Haryati S., Sudarsono A., Suryana E., 2015).

The development of knowledge and technology has resulted in a high level of competition in the world of education which makes every educational institution competing in providing quality service to all the infrastructure provided, one of which is the library.

STMIK Pelita Nusantara is an educational foundation that always follows the progress of education, the development of knowledge and technology that can increase awareness of service users towards the quality of services provided by an agency. The quality of services in the world of education provided by an institution becomes very important because it can help agencies in providing quality services that are as expected, especially for library visitors at STMIK Pelita Nusantara. In this study library visitors at STMIK Pelita Nusantara are samples that will provide an assessment of the quality of services provided, visitors will compare the expected services with the services they receive. The quality of service in accordance with what is expected by visitors to the library will make it an attraction for visitors to become fond of visiting the library reading books that will help academically visitors who are mostly students.

Based on the problems encountered and the results of previous studies that the author has explained, the authors are interested in conducting a study to measure the satisfaction of library visitors using the C4.5 algorithm. With the Special Purpose is to create an information system that can map the level of satisfaction of library users at STMIK Pelita Nusantara in order to provide

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

information to the Management of STMIK Pelita Nusantara in Improving Excellent Service for library users who can support the quality of Graduates.

2. Theoretical Basis

A. *Knowledge Discovery in Database (*KDD*)*

Apart from the understanding that has been explained in the background, data mining is also called Knowledge Discovery in Database (KDD) which is defined as the extraction of implicit and unknown potential information from a set of data. The Knowlegde Discovery in Database process involves the results of the data mining process (the process of extracting the tendency of a data pattern), then converts the results accurately into information that is easily understood (Tampubolon K., Saragih H. and Reza B., 2013).

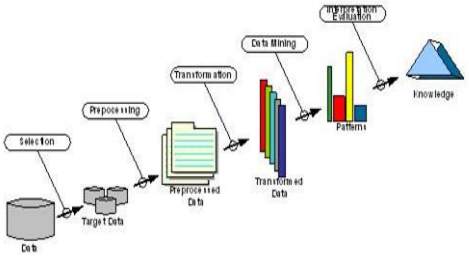

Fig 1. *Knowledge Database Discovery* (KDD)

B. Data Mining

Data mining is the process of extracting data into information that has not been conveyed before, with the right technique the data mining process will provide optimal results (Abdillah G., Putra F. A. and Renaldi F., 2016).

C. Algoritma C4.5

C4.5 algorithm or decision tree is a method commonly used to classify data mining. As explained earlier, classification is a technique of finding a collection of patterns or functions that describe and separate one class of data from another to declare the object into a particular category by looking at the behavior and attributes of the group that has been defined. This method is popular because it is able to do the classification while showing the relationship between attributes. Many algorithms can be used to build a decision tree, one of which is the C4.5 algorithm. C4.5 algorithm can handle numeric and discrete data. C4.5 algorithm uses a gain ratio. Before calculating the acquisition ratio, it is necessary to calculate the value of information in units of bits from a collection of objects, namely by using the concept of entropy (Supriyanti W., Kusrini and Amborowati A., 2016).

There is also a mention that the C4.5 algorithm constructs decision trees from training data, in the form of cases or records (tuples) in the database (Putra D. W. T., 2016).

1) Concept of Entropy

Entropy (S), which is the estimated number of bits needed to be able to extract a class (+ or -) from a random amount of data in the S sample space. Entropy can be said as the need for bits to express a class. The smaller the value of entropy, the more entropy is used in extracting a class. Entropy is used to measure the authenticity of S. The amount of Entropy in the sample space S is defined by:

$$\text{Enthropy}(S) = \sum_{i=1}^{n} - pi * \log_2 pi \qquad (1)$$

Where :

S : case set
A : the features
n : number of partitions S
pi : the proportion of Si to S

2) The concept of gain

Gain (S, A) is the acquisition of information from attribute A relative to the output data S. The acquisition of information is obtained from the output data or dependent variable S which is grouped based on attribute A, denoted by gain (S, A).

$$\text{Gain}(S,A) = \text{Enthropy}(S) - \sum_{i=1}^{n} - \frac{|Si|}{|S|} * \text{Enthopy}(Si) \qquad (2)$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Dengan :
S : Case set
A : Atribute
n : Number of attribute attributes A
|Si| : Number of cases on partition i
|S| : Number of cases in S

3. Discussion

A. Application of C4.5 Algorithm

The design of the system will be used in determining the conditions in measuring the level of library visitor satisfaction. The initial data to be divided in the form of numerical or non-numeric class to facilitate further analysis. After all the data to be processed is divided by class, then the classification process will be carried out by forming a decision tree as the output.

The decision making process to determine the conditions for measuring student satisfaction levels is as follows:

1) *Tangibles*
2) *Reability*
3) *Responsiveness*
4) *Assurance*
5) *Empathy*

The categories that will become the decision are Satisfied and Dissatisfied.

The process of processing data mining methods using the C.45 algorithm to produce rules of the conditions in measuring the level of student satisfaction can be seen in Table 1:

Table 1.
Final Data Questionnaire Results Measuring Satisfaction Level of Library Visitors

No	Tangiable	Reability	Responsivnes	Assurance	Empathy	Hasil
1	Cukup	Tinggi	Tinggi	Cukup	Tinggi	Puas
2	Rendah	Tinggi	Cukup	Cukup	Rendah	Tidak Puas
3	Tinggi	Tinggi	Rendah	Cukup	Tinggi	Puas
4	Rendah	Rendah	Rendah	Cukup	Cukup	Tidak Puas
5	Cukup	Cukup	Cukup	Cukup	Cukup	Puas
6	Cukup	Cukup	Cukup	Cukup	Rendah	Puas
7	Rendah	Cukup	Tinggi	Cukup	Tinggi	Puas
8	Rendah	Cukup	Tinggi	Tinggi	Tinggi	Puas
9	Tinggi	Cukup	Cukup	Tinggi	Tinggi	Puas
10	Cukup	Cukup	Cukup	Tinggi	Cukup	Puas
11	Cukup	Cukup	Cukup	Rendah	Rendah	Puas
12	Rendah	Cukup	Cukup	Rendah	Rendah	Tidak Puas
13	Tinggi	Cukup	Tinggi	Cukup	Cukup	Puas
14	Cukup	Cukup	Tinggi	Cukup	Tinggi	Puas
15	Cukup	Cukup	Tinggi	Cukup	Cukup	Puas
16	Tinggi	Rendah	Tinggi	Tinggi	Tinggi	Puas
17	Cukup	Rendah	Tinggi	Cukup	Cukup	Puas
18	Cukup	Rendah	Tinggi	Cukup	Tinggi	Puas

B. Classification With C4.5 Algorithm

The case listed will make a decision tree to identify the conditions in measuring the level of library visitor satisfaction based on the results of the library visitor questionnaire at STMIK Pelita Nusantara Medan that has been distributed by looking at the questionnaire results in measuring student satisfaction levels (Tangibles, Reability, Responsiveness, Assurance, Empathy). To choose

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

an attribute as the root, based on the highest gain value of the existing attributes. To calculate the gain used formula (2), while to calculate the value of entropy can be seen in the formula (1).

In making a decision tree, what must be done is to count the number of cases, the number of cases for a "Satisfied" decision (S1), the number of cases for a "Dissatisfied" (S2) decision and cases divided by the attributes of Tangibles, Reability, Responsiveness, Assurance, Empathy , then we calculate the gain for each attribute. The steps for making a decision tree are:

Determine the attribute as the root and calculate the attribute gain information value.

To choose an attribute as the root, based on the highest gain value of the existing attributes. Entropy value is needed to determine the highest gain.

Calculate the Entropy Value for each attribute:

Entropy (Total) with the following formula:

Entropy (*total*) = $-\sum_{i=1}^{n} -\mathrm{P}\, i \times \log_2 \mathrm{P}i$

$$Entropy(total) = \left(-\frac{15}{18} * log_2\left(\frac{15}{18}\right)\right) + \left(-\frac{3}{18} * log_2\left(\frac{3}{18}\right)\right) = 0.65002$$

Entropy (total) is counting the total value of decisions received (15) and not received (3), while 18 is the total number of cases.

1. *Tangibles*

   To calculate Tangibles entropy seen from table 4.6, Tangibles values received which have a High value of 4 cases, Tangibles with a sufficient value of 9 cases and Tangibles with a low value of 5 cases, the following is the entropy of each case:

   $$Entropy(T) = \left(-\frac{4}{4} * log_2\left(\frac{4}{4}\right)\right) + \left(-\frac{0}{4} * log_2\left(\frac{4}{0}\right)\right) = 0$$

   $$Entropy(C) = \left(-\frac{9}{9} * log_2\left(\frac{9}{9}\right)\right) + \left(-\frac{0}{9} * log_2\left(\frac{0}{9}\right)\right) = 0$$

   $$Entropy(R) = \left(-\frac{2}{5} * log_2\left(\frac{2}{5}\right)\right) + \left(-\frac{3}{5} * log_2\left(\frac{3}{5}\right)\right) = 0.97095$$

   For cases with other variables, do the same thing.

2. *Reability*

   $$Entropy(T) = \left(-\frac{2}{3} * log_2\left(\frac{2}{3}\right)\right) + \left(-\frac{1}{3} * log_2\left(\frac{1}{3}\right)\right) = 0.91829$$

   $$Entropy(C) = \left(-\frac{10}{11} * log_2\left(\frac{10}{11}\right)\right) + \left(-\frac{1}{11} * log_2\left(\frac{1}{11}\right)\right) = 0.13752$$

   $$Entropy(R) = \left(-\frac{3}{4} * log_2\left(\frac{3}{4}\right)\right) + \left(-\frac{1}{4} * log_2\left(\frac{1}{4}\right)\right) = 0.81128$$

3. *Responsiveness*

   $$Entropy(T) = \left(-\frac{9}{9} * log_2\left(\frac{9}{9}\right)\right) + \left(-\frac{0}{9} * log_2\left(\frac{0}{9}\right)\right) = 0$$

   $$Entropy(C) = \left(-\frac{5}{7} * log_2\left(\frac{5}{7}\right)\right) + \left(-\frac{2}{7} * log_2\left(\frac{2}{7}\right)\right) = 0.86312$$

   $$Entropy(R) = \left(-\frac{1}{2} * log_2\left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} * log_2\left(\frac{1}{2}\right)\right) = 1$$

4. *Assurance*

   $$Entropy(T) = \left(-\frac{4}{4} * log_2\left(\frac{4}{4}\right)\right) + \left(-\frac{0}{4} * log_2\left(\frac{0}{4}\right)\right) = 0$$

   $$Entropy(C) = \left(-\frac{10}{12} * log_2\left(\frac{10}{12}\right)\right) + \left(-\frac{2}{12} * log_2\left(\frac{2}{12}\right)\right) = 0.65002$$

   $$Entropy(R) = \left(-\frac{1}{2} * log_2\left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} * log_2\left(\frac{1}{2}\right)\right) = 1$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

5. *Empathy*

$$Entropy(T) = \left(-\frac{8}{8} * log_2\left(\frac{8}{8}\right)\right) + \left(-\frac{0}{8} * log_2\left(\frac{0}{8}\right)\right) = 0$$

$$Entropy(C) = \left(-\frac{5}{6} * log_2\left(\frac{5}{6}\right)\right) + \left(-\frac{1}{6} * log_2\left(\frac{1}{6}\right)\right) = 0.65002$$

$$Entropy(R) = \left(-\frac{2}{4} * log_2\left(\frac{2}{4}\right)\right) + \left(-\frac{2}{4} * log_2\left(\frac{2}{4}\right)\right) = 1$$

Menghitung nilai*gain* tiap-tiap atribut:

1. *Gain* (Total, *Tengibles*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Tengibles_i|}{|Total|} * Entropy(Tengibles_i)$$

$$= 0.65002 - \left(\left(\frac{4}{18} * 0\right) + \left(\frac{9}{18} * 0\right) + \left(\frac{5}{18} * 0.97095\right)\right) = 0.38031$$

2. *Gain* (Total, *Reability*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Reability_i|}{|Total|} * Entropy(Reability_i)$$

$$= 0.65002 - \left(\left(\frac{3}{18} * 0.91829\right) + \left(\frac{11}{18} * 0.13752\right) + \left(\frac{4}{18} * 0.81128\right)\right) = 0.23265$$

3. *Gain* (Total, *Responsiveness*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Responsiveness_i|}{|Total|} * Entropy\ (Responsiveness_i)$$

$$= 0.65002 - \left(\left(\frac{9}{18} * 0\right) + \left(\frac{7}{18} * 0.86312\right) + \left(\frac{2}{18} * 1\right)\right) = 0.20325$$

4. *Gain* (Total, *Assurance*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Assurance_i|}{|Total|} * Entropy(Assurance_i)$$

$$= 0.65002 - \left(\left(\frac{4}{18} * 0\right) + \left(\frac{12}{18} * 0.65002\right) + \left(\frac{2}{18} * 1\right)\right) = 0.10556$$

5. *Gain* (Total, *Empathy*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Empathy_i|}{|Total|} * Entropy(Empathy_i)$$

$$= 0.65002 - \left(\left(\frac{8}{18} * 0\right) + \left(\frac{6}{18} * 0.65002\right) + \left(\frac{4}{18} * 1\right)\right) = 1$$

After all entropy values and the gain of each attribute are calculated, the results of these calculations are entered in table 2.

Table 2
Calculation of Node 1

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Node	Jumlah Kasus(S)			Puas	Tidak Puas	Entropy	Info Gain
1	Total		18	15	3	0.65002	
	Tangible						0.380314
		Tinggi	4	4	0	0	
		Cukup	9	9	0	0	
		Rendah	5	2	3	0.97095	
	Reability						0.232652
		Tinggi	3	2	1	0.91829	
		Cukup	11	10	1	0.13752	
		Rendah	4	3	1	0.81128	
	Responsivnes						0.203253
		Tinggi	9	9	0	0	
		Cukup	7	5	2	0.86312	
		Rendah	2	1	1	1	
	Assurance						0.105563
		Tinggi	4	4	0	0	
		Cukup	12	10	2	0.65002	
		Rendah	2	1	1	1	
	Empathy						0.211126
		Tinggi	8	8	0	0	
		Cukup	6	5	1	0.65002	
		Rendah	4	2	2	1	

From the calculations in the table above it can be seen that the attribute with the highest gain is Responsiveness with a gain value of 0.380314 then the attribute is used as the root node or node 1, where the attribute value is sufficient and the attribute value is low which still has to be recalculated, the decision tree generated from the node 1 can be seen in Figure 2 below.

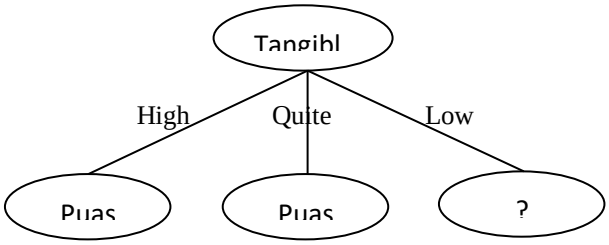

Fig 2. Decision Tree Results of Calculation Node 1

The next step is completed to calculate Node 1 and Node 1.1 as roots for the sufficient and low values of the Tangible attribute, the same as the previous method above by calculating the entropy value of the remaining attributes namely Reability, Responsiveness, Assurance and Empathy after the entropy is calculated, then calculate the gain for each attribute.

Entropy (Low Tangible Value), with the following formula:

Entropy (*total*) = $-\sum_{i=1}^{n} -\mathrm{P}\,i \times \log_2 \mathrm{P}i$

$$= \left(-\frac{2}{5} * log_2\left(\frac{2}{5}\right)\right) + \left(-\frac{3}{5} * log_2\left(\frac{3}{5}\right)\right) = 0.970951$$

1) Atribute *Reability*

$$Entropy(T) = \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) = 0$$

$$Entropy(C) = \left(-\frac{2}{3} * log_2\left(\frac{2}{3}\right)\right) + \left(-\frac{1}{3} * log_2\left(\frac{1}{3}\right)\right) = 0.918296$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

$$Entropy(R) = \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) = 0$$

2) Atribute *Responsivenes*

$$Entropy\ (T) = \left(-\frac{2}{2} * log_2\left(\frac{2}{2}\right)\right) + \left(-\frac{0}{2} * log_2\left(\frac{0}{2}\right)\right) = 0$$

$$Entropy\ (C) = \left(-\frac{0}{2} * log_2\left(\frac{0}{2}\right)\right) + \left(-\frac{2}{2} * log_2\left(\frac{2}{2}\right)\right) = 0$$

$$Entropy\ (R) = \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) = 0$$

3) Atribute *Assurance*

$$Entropy(T) = \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) = 0$$

$$Entropy(C) = \left(-\frac{1}{3} * log_2\left(\frac{1}{3}\right)\right) + \left(-\frac{2}{3} * log_2\left(\frac{2}{3}\right)\right) = 0.918296$$

$$Entropy(R) = \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) = 0$$

4) Atribute *Empathy*

$$Entropy(T) = \left(-\frac{2}{2} * log_2\left(\frac{2}{2}\right)\right) + \left(-\frac{0}{2} * log_2\left(\frac{0}{2}\right)\right) = 0$$

$$Entropy(R) = \left(-\frac{0}{1} * log_2\left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * log_2\left(\frac{1}{1}\right)\right) = 0$$

$$Entropy(C) = \left(-\frac{0}{2} * log_2\left(\frac{0}{2}\right)\right) + \left(-\frac{2}{2} * log_2\left(\frac{2}{2}\right)\right) = 0$$

Calculates the value of each attribute:

1) *Gain* (Total, *Reability*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Tes\ Reability_i|}{|Total|} * Entropy(Reability_i)$$

$$= 0.970951 - \left(\left(\frac{1}{5} * 0\right) + \left(\frac{3}{5} * 0.918296\right) + \left(\frac{1}{5} * 0\right)\right) = 0.419973$$

2) *Gain* (Total, *Responsivnes*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Responsivenes_i|}{|Total|} * Entropy(Responsivenes_i)$$

$$= 0.970951 - \left(\left(\frac{2}{5} * 0\right) + \left(\frac{2}{5} * 0\right) + \left(\frac{1}{5} * 0\right)\right) = 0.970951$$

3) *Gain* (Total *Assurance*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Assurance_i|}{|Total|} * Entropy(Assurance_i)$$

$$= 0.970951 - \left(\left(\frac{1}{5} * 0\right) + \left(\frac{3}{5} * 0.918296\right) + \left(\frac{1}{5} * 0\right)\right) = 0.419973$$

4) *Gain* (Total, *Empathy*)

$$= Entropy(S) - \sum_{1=1}^{n} \frac{|Empathy_i|}{|Total|} * Entropy(Empathy_i)$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

$$= 0.970951 - \left(\left(\frac{2}{5} * 0\right) + \left(\frac{1}{5} * 0\right) + \left(\frac{2}{5} * 0\right)\right) = 0.970951$$

Furthermore, after the entropy and gain values are calculated, the results of these calculations are included in table 3.

Table 3.
Calculation of Nodes 1.1

Node	Jumlah Kasus(S)			Puas	Tidak Puas	Entropy	Info Gain
1	Total		5	2	3	0.970951	
	Reability						0.419973
		Tinggi	1	0	1	0	
		Cukup	3	2	1	0.918296	
		Rendah	1	0	1	0	
	Responsivnes						0.970951
		Tinggi	2	2	0	0	
		Cukup	2	0	2	0	
		Rendah	1	0	1	0	
	Assurance						0.419973
		Tinggi	1	1	0	0	
		Cukup	3	1	2	0.918296	
		Rendah	1	0	1	0	
	Empathy						0.970951
		Tinggi	2	2	0	0	
		Cukup	1	0	1	0	
		Rendah	2	0	2	0	

From the table above it can be seen that the highest gain attribute used as Node 1.1 is the Reability attribute, with a gain value of 0.970951. Thus it can be interpreted that Assurance can be a root node 1.1. The decision tree formed can be seen in Figure 3.

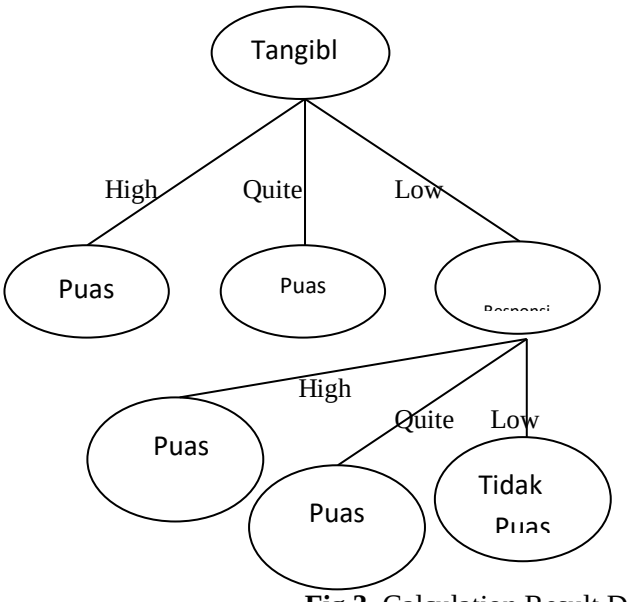

Fig 3. Calculation Result Decision Tree Node 1.1

Thus it can be interpreted that Responsivnes can be a root node 1.1. The decision tree formed can be seen in Figure 3.

The rules that are formed based on the last decision tree as in Figure 4.5 above are as follows: berdasarkan pohon keputusan terakhir seperti pada gambar 4.5 diatas adalah sebagai berikut :

1. IF *Tangible* = Tinggi THEN Keptusan = Puas
2. IF *Tangible* = Cukup THEN Keptusan = Puas

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

3. IF *Tangible* = Rendah AND *Responsiveness* = Tinggi THEN Keputusan = Puas
4. IF *Tangible* = Rendah AND *Responsiveness* = Cukup THEN Keputusan = Tidak Puas
5. IF *Tangible* = Rendah AND *Responsiveness* = Rendah THEN Keputusan = Tidak Puas

C. Implementation with RapidMiner

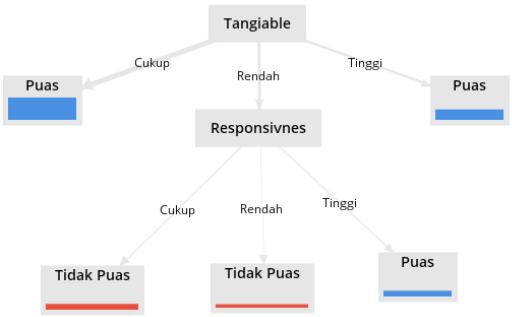

Fig 4. Displaying Test Results with RapidMiner Study 9.7

4. Conclusion

Based on the results, several conclusions can be drawn as follows:

1) Decision Tree model C4.5 algorithm can be used to classify variables in the negativity level of library user satisfaction that can be used as a basis in making a decision support system in improving services to library users.
2) Implementation results have been tested using one of the Data Mining applications with the C4.5 algorithm, RapidMiner Studio 9.7 and based on the test results obtained the same classification results.

5. References

[1] E. Khoirsta, Afrinda, Yulianto and M. K. Mawardi, "(Survey Pada Pelanggan Fedex Express Surabaya)," *J. Adm. Bisnisdministrasi Bisnis*, vol. 25, no. 2, pp. 1–7, 2015.

[2] D. Dicky Nofriansyah, "Penerapan Data Mining dengan Algoritma Naive Bayes Clasifier untuk Mengetahui Minat Beli Pelanggan terhadap Kartu Internet XL (Studi Kasus di," *Saintikom*, vol. 15, no. 1978–6603, pp. 81–92, 2016.

[3] D. W. T. Putra, "Algoritma C4.5 untuk Menentukan Tingkat Kelayakan Motor Bekas yang Akan Dijual," *J. TEKNOIF*, vol. 4, no. 1, pp. 16–22, 2016.

[4] F. Nasari and S. Darma, "Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru," *Semin. Nas. Teknol. Inf. dan Multimed. 2015*, pp. 73–78, 2015.

[5] "International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS) TALENT KNOWLEDGE ACQUISITION USING C4 . 5 CLASSIFICATION ALGORITHM," pp. 406–410, 2013.

[6] K. Tampubolon, H. Saragih, B. Reza, K. Epicentrum, A. Asosiasi, and A. Apriori, "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan," pp. 93–106, 2013.

[7] J. Eska, "Penerapan Data Mining Untuk Prekdiksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksiaran," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, pp. 9–13, 2016.

[8] I. Rahmayuni, "Perbandingan performansi algoritma c4.5 dan cart dalam klasifiksi data nilai mahasiswa prodi teknik komputer politeknik negeri padang," *Teknoif*, vol. 2, no. 1, pp. 40–46, 2014.

[9] F. Riandari and A. Simangunsong, "Penerapan Algoritma C4.5 Untuk Mengukur Tingkat Kepuasan Mahasiswa," *Terakreditasi DIKTI*, vol. 3, no. 2, pp. 1–7, 2019.