

CHAPTER 3

RESEARCH METHOD

3.1 Research Design

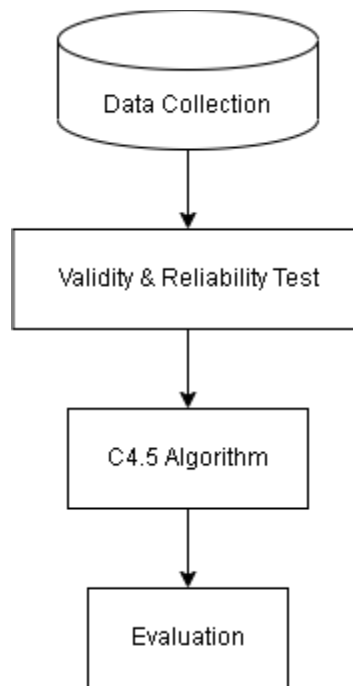


Figure 3.1 Research Design

Based on figure 2.1, this research begins by collecting data by distributing questionnaires to respondent. The collected data will be tested for feasibility as a research instrument with validity and reliability test. Then the C4.5 algorithm is calculated on the data set to generate a decision tree. Evaluation of the C4.5 model accuracy using Split Validation.

3.2 Data Collection

The data collection was carried out by distributing questionnaires, the variable of the questionnaire can be seen in Table 3.1. The sample of this study is active students in the Department of Informatics at Universitas Muhammadiyah Kalimantan Timur class 2019. Scale measuring tool using a Likert scale which includes a score of 1 (very dissatisfied), a score of 2 (not satisfied), a score of 3

(satisfied), and a score of 4 (very satisfied) for each questions about learning performance by indicator point.

Table 3.1 Questions on The Questionnaire According to The Indicator

NO	QUESTIONS	INDICATOR
1.	Textbooks, workbooks and/or handouts are well provided	Tangible
2.	Bibliography needed for the learning process of the courses given by the lecturer	
3.	Using an LCD projector in learning makes it easier for me to understand what is being taught	
4.	Students easily get reading material or references to achieve the target of lectures	
5.	The campus library provides adequate references in the form of books, journals, and reading materials	
6.	Classrooms are arranged cleanly and neatly when they are going to be used	
7.	Classrooms are comfortable to use	
8.	Facilities that support the learning process are available in full in the classroom (class computers, LCD projectors, whiteboards) and replacements are available in case of damage.	
9.	Laboratory facilities meet the needs of students in the learning process	
10.	There are internet facilities that are easily accessible in the campus environment	
11.	Availability of adequate and clean toilet facilities	
12.	It is easy to find reading materials to complete this course	
13.	I understand very well the conditions that are desired from us in the learning process of this course	Responsiveness
14.	I can clearly see how the parts of this lecture are put together well to achieve the desired target	
15.	I understand how our study activities are assessed	
16.	At the beginning of the lecture, the lecturer conveys about the abilities that students get after attending this lecture	Assurance
17.	At the beginning of the lecture the lecturer explained well about the assessment system for students	
18.	Lecturers start and end lectures according to learning hours	
19.	Lecturers prepare lecture materials well	
20.	Lecturers explain and facilitate learning activities well	

NO	QUESTIONS	INDICATOR
21.	Lecturers give group assignments as effectiveness in learning	
22.	Lecturers give assignments at the end of each lecture time	
23.	The results of the correction of assignments are immediately given by the lecturer	
24.	The method of assessment in lectures is in accordance with what was agreed at the beginning of the lecture	
25.	Assessment takes into account the results of individual work and group work	
26.	Assessment is considered through the activeness of students during learning	
27.	Lecturers inspire and motivate students	Empathy
28.	Students get sufficient guidance during the active period of study from academic supervisors	
29.	There is good academic guidance on which course choices suit my needs and interests	
30.	The subjects studied get a good understanding of the field	Reliability
31.	I have learned to think critically as a result of learning during face-to-face lectures	
32.	I have learned to present ideas in a clear way as a result of learning activities	
33.	I have developed communication skills as a result of learning activities during face-to-face lectures	
34.	I have developed my ability to work in groups or teams as a result of the work I do in learning during face-to-face lectures	

3.3 Validity and Reliability Test

a. Validity Test

Validity test is used to measure whether or not the used variable is valid. To measure the validity of a variable using Equation 3.1. After getting the r_{xy} value for each variables, then r_{xy} compared to the r table. If r_{xy} is smaller than r table then the variable item is stated invalid and if r_{xy} is greater than r table then the variable is stated valid.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}} \quad (3.1)$$

where r_{xy} is correlation coefficient X and Y, N is number of subjects, $\sum XY$ is X and Y multiplication score, $\sum X$ is amount of X, $\sum X^2$ is sum of X squares, $\sum Y$ is amount Y, and $\sum Y^2$ is sum of Y squares.

b. Reliability Test

Reliability test is used to shows the measuring instrument can be trusted or reliable. To measure a questionnaire is reliable or not using, calculating the reliability using Equation 3.2. To determine the level of reliability of the questionnaire using Guilford Reliability Coefficient Category, can be seen at Table 3.2.

$$r_{11} = \left(\frac{K}{k-1} \right) \left(1 - \frac{\sum \sigma_b^2}{\sigma_1^2} \right) \quad (3.2)$$

where r_{11} is instrument reliability, K is many questions, $\sum \sigma_b^2$ is varians items summary, and σ_1^2 is total varians.

Table 3.2 Guilford Reliability Coefficient Category

Scale	Coefficient
$0,80 \leq r_{11} \leq 1,00$	Very High
$0,60 \leq r_{11} \leq 0,80$	High
$0,40 \leq r_{11} \leq 0,60$	Medium
$0,20 \leq r_{11} \leq 0,40$	Low
$0,00 \leq r_{11} \leq 0,20$	Very Low

3.4 C4.5 Algorithm

The C4.5 algorithm uses entropy and gain for the selection of attributes to generate the decision tree. The biggest gain value will be the root node or the first node. The decision tree participation process stops when all the branches in node N are in the same class, there are no attributes in the branch that can be subdivided and there are no empty branches. To calculate the entropy using Equation 3.3 and to calculate gain using Equation 3.4.

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (3.3)$$

where S is set of cases, n is number of partitions s, and p_i os proportion of S_i to S.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S) \quad (3.4)$$

where S is set of cases, A is attribute, n is number of partitions, $|S_i|$ is number of cases on partition i , and $|S|$ is number of cases in S .

3.5 Evaluation

Evaluation is carried out to analyze and measure the accuracy of C4.5 algorithm model using Split Validation. Split Validation is a technique that random the data and divides data into training data and testing data with a ratio of 7:3.