

CHAPTER III RESEARCH METHODOLOGY

3.1 Introduction

In this chapter, there will be a research method used in this research. In this section, the content-based filtering testing method (X1), cluster analysis recommendations (X2) and the collaborative filtering recommender system (X3) on online consumer behavior (Y) are included. The information about the methods used by the researcher to analyze the relationship between the dependent variable and the independent variable, the research methodology techniques, the evaluation and interpretation of the data collected by the researcher are described in detail.

3.2 Research Design

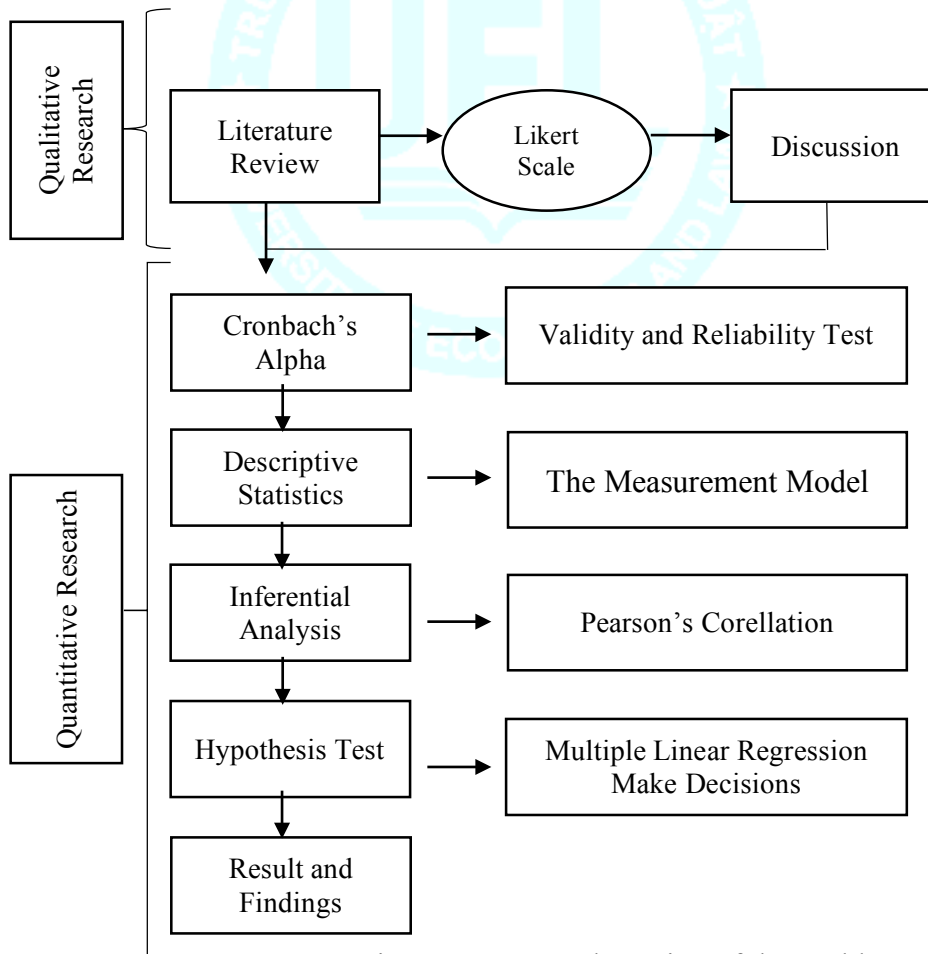


Figure 1 : Research Design of the Problem

3.3 Research Types

This research uses mix method which is quantitative dan qualitative research. Research with a quantitative approach emphasizes its analysis on numerical data (numbers) processed by statistical methods. Basically, the quantitative approach is carried out in inferential research (in the context of testing the hypothesis) and rests the conclusion of the results on a probability of null hypothesis rejection (H_0). qualitative research, including observations, textual or visual analysis (eg from books or videos) and interviews (individual or group). In this research qualitative methods use to make the questionnaire and collect the data and for the quantitative is for count and processing data.

3.4 Data Source

The data generated by the researcher is the result of the processing process during the research. The type of data used in this research is primary data. According to Hasan (2002: 82) primary data is data obtained or collected directly in the field by the person conducting the research or those concerned who need it.

3.5 Data Collection Method

The research data were obtained using data collection instruments. The data instrument used in this study was the Likert scale. Likert scale is a scale that contains a statement about the object to be studied. This scale has two attitude statements, namely favorabel (supporting or in favor of the subject) and unfavorable (not supporting the subject) (Azwar, 1998: 98).

3.5.1 Questionnaire

The questionnaire is a data collection technique by providing a written statement to the respondent. According to Sugiyono (2013: 199) a questionnaire is a data collection technique that is done by giving a set of questions or written statements to respondents to be answered.

3.5.2 Questionnaire Design

Table 1 : Questionnaire Design

SA	A	N	D	TD
(5)	(4)	(3)	(2)	(1)

3.6 Population And Sampling

Hair et al., (2010) suggested, for models of five or fewer structures, that the minimum sample size is 100. Such a sample size will contribute to the interpretation of the relationship and the significance of the study variables being identified. Therefore, the researcher decides to select 200 respondents for the research that the researcher considers this sample to be significant to achieve accurate results and findings.

Table 2 : Sampling Techniques and Determination of Sample Size of Population Applied

Source : Krejcie & Morgan, 1970 (Semantic Scholar)

Nominal Level—Parallel Design with Raw Data

$\sigma\%$	θ' (Power 80%)				θ' (Power 90%)			
	0%	5%	10%	15%	0%	5%	10%	15%
10	10	12	26	100	14	18	36	138
12	14	18	38	144	18	24	52	198
14	18	24	50	196	24	32	68	270
16	24	30	64	254	30	40	90	352
18	30	38	82	322	38	52	112	446
20	36	46	100	398	46	62	138	550
22	44	54	122	480	54	76	168	664
24	52	64	144	572	64	90	198	790
26	60	76	168	670	76	104	234	928
28	68	88	196	778	86	122	270	1076
30	78	100	224	892	100	138	310	1234
32	90	114	254	1014	112	158	352	1404
34	100	128	288	1146	126	178	398	1586
36	112	144	322	1284	142	198	446	1778
38	126	160	358	1430	158	222	496	1980
40	138	178	398	1584	176	246	550	2194

3.6.1 Population

According to Nawawi (in Utari 2005: 13), population is the total object of research which can consist of humans, objects, animals, plants, symptoms, test scores or events, as a source of data that has certain characteristics in a study. The population in this study are consumers around the world who know about advertising and use technology in their daily lives.

3.6.2 Sample

Samples can be interpreted as part of the population that is the actual source of data in a study. That is, the sample is a part of the population to represent the entire population, Nawawi (in Utari 2005: 13). The number of samples used in this study were 100 respondents. The sampling in this research was *probability sampling*. Sugiyono (2014), Probability sampling is a sampling technique that provides equal opportunities for each element (member) of the population to be selected as a sample. Then *using simple random sampling*. According to Kerlinger (2006: 188), Simple random sampling is a method of drawing from a population or universe in a certain way so that each member of the population or universe has an equal chance of being selected or taken. Hair et al (2010) suggest that the number of research samples with unknown population size means that the sample size is at least 5 times the number of statements. Hair et all also suggested that a suitable sample size ranged from 100-200 respondents. Then the sample calculation results are obtained as follows:

The questions of the questionnaire	= 25 questions
Respondents	= 25 x 8
	= 200 respondents

3.7 Reseach Variable

According Dr. Soekidjo Notoatmojo (2002), namely :

- a. Variables contain the sense of size or characteristics possessed by members in a group that are different from what other groups have.
- b. A variable is something that is used as a trait, trait, or trait obtained from research on certain concepts. For example, education, age, genes,

occupation, knowledge, and so on.

3.7.1 Independent

Variable Independent variable is a variable that is not dependent on other variables. In this study the variables are :

X1 : **Content-based** Filtering

X2 : Recommendation Cluster Analysis

X3 : Collaborative Filtering Recommender System

3.7.2 Dependent

The dependent variable is a variable that depends on other variables. In this study the variable is :

Y : Online Consumer Satisfaction.

3.8 Data Analysis

The analysis was carried out by examining the relationship between the dependent variable and several independent variable. This study uses Multiple Linear Regression (MLR) because there is one the dependent variable and independent four variables. This technique is used to find a relationship between a dependent variable and one or more independent variables (Sugiyono, 2009: 275). Multiple linear regression (MLR) is as statistical method that us widely used in HKSA studies, buy the results of analysis with this method have low accuracy (Larif et al., 2013). Because in MLR the relationship between structure and activity is described as a linear-relationship, Abstract. The selection of statistical techniques that will be used is adjusted to the predetermined research design, namely looking for the impact of the variable X1 : Content-based Filtering, variable X2: Recommendation Cluster Analysis and variable X3: Collaborative Filtering Recommender System on Online Consumer Satisfaction (Y). Analysis of this data using the computerized help of the SPSS (Statistical Product and Service Solutions) 20 For Mac program.

Multiple regression analysis is used to analyze the magnitude of the relationship and the effect of the independent (free) variable which is more than two. The multiple regression analysis equation can be formulated as follows:

$$Y^1 = a + b_1X_1 + b_2X_2 + b_3X_3$$

Soucre : Sugiyono (2015:192)

3.8.1 Pilot Test

The pilot test was used to test the reliability and validity of the research instrument. Before the questionnaire was distributed to respondents, the questionnaire was tested first on 30 individual consumer e-commerce users in Vietnam (Ho Chi Minh City) and Indonesia (Bandar Lampung).

3.8.2 Validity

Validity comes from the word validity which means the extent to which the accuracy and accuracy of a measuring instrument in carrying out its measure function (Azwar 1986). In addition, validity is a measure that indicates that the measured variable is really the variable that the researcher wants to study (Cooper and Schindler, in Zulganef, 2006). Meanwhile, according to Sugiharto and Sitinjak (2006), validity is related to a variable measuring what should be measured. The validity of the research states the degree of accuracy of the research measuring instrument to the actual content being measured. The validity test is a test used to show the extent to which measuring instruments are used in measuring what is being measured. Ghozali (2009) states that the validity test is used to measure whether a questionnaire is valid or not. A questionnaire is said to be valid if the questions on the questionnaire are able to reveal something that will be measured by the questionnaire. The validity test is used to find out how accurate an item is in measuring what it wants to measure. A question item is said to be valid if it is a value the correlation is greater than the correlation value with the total score (Priyatno, 2018). Valid conditions or not can be determined by the following criteria:

Information :

X : Item Score

Y : Total score

N : Number of respondents to be tested

r : Product moment correlation

Testing procedure:

1. H₀: valid data
Ha: invalid data
2. H₀: what if $r \text{ count} > r \text{ table}$ then the instrument is declared valid
Ha: what if $r \text{ count} < r \text{ table}$ then the instrument is declared invalid.
3. Testing the validity instrument was carried out using the SPSS 20.0 program.
4. Explanation of point 1 and point 2 by comparing $r \text{ count}$ and $r \text{ table}$, it can be concluded that the instrument is valid or invalid.

3.8.3 Reliability

The definition of reliability (reliability) is the consistency of measurement (Walizer, 1987). Sugiharto and Situnjak (2006) state that reliability refers to an understanding that the instruments used in research to obtain the information used can be trusted as a data collection tool and are able to reveal actual information in the field. Ghozali (2009) states that reliability is a tool for measuring a questionnaire which is an indicator of variables or constructs. A questionnaire is said to be reliable or reliable if a person's answer to a statement is consistent or stable over time. The reliability of a test refers to the degree of stability, consistency, predictive power, and accuracy. Measurements that have high reliability are measurements that can produce reliable data. According to Sumadi Suryabrata (2004: 28) reliability shows to what extent the measurement results with these tools can be trusted. The measurement results must be reliable in the sense that they must have a level of consistency and stability. According to Sumadi Suryabrata (2004: 28) reliability shows to what extent the measurement results with these tools can be trusted.

3.8.4 Descriptive Statistics

Makhfudli et al., (2019) stated that in an analysis of the data, a descriptive statistical approach was used to produce the results of the study in frequency, percentage, mean, ranges and standard deviation forms. This method allowed the study subject characteristics to be summarized based on the selected variables. Descriptive research is one type of research whose

purpose is to present a complete picture of a social setting or is intended for exploration and clarification of a social phenomenon or reality. Do this by describing a number of variables with respect to the problem and the unit under study between the phenomena being tested. According to Nazir in his research method, descriptive method is a method in examining the status of human groups, a subject, a set of conditions, a system of thought or a class of events in the present.

3.8.5 Hypothesis Test (Multiple Linear Regression)

According to Suharyadi & Purwanto (2011: 86) “The statistical test is a value obtained from the sample and determine whether to accept or reject the hypothesis.” To find out how much influence the cost of prevention and appraisal costs on the product, the test is carried out using regression analysis. According to V. Wiratna (2015: 111) “Regression aims to test between one variable and another.”

Information :

Y : Dependent Variable (Online Consumer Satisfaction).

a : Constant / Y value, if X = 0

$b_1 b_2 b_3$: The regression direction coefficient, which states the change in Y value if there is a change in X value.

X_1 : Independent Variable (Content-based Filtering).

X_2 : Independent Variable (Recommendation Cluster Analysis).

X_3 : Independent Variable (Collaborative Filtering Recommender System).

Meanwhile, the values a and b can be determined by the equation as follows:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

Information:

a : Constant

b : Regression coefficient

X : Independent variable
Y : Dependent variable
n : Number of samples

3.9 Summary

Chapter three provides an overview of the research methodology used in this study. In order for researchers to get accurate findings, the right method is very important. Descriptive research was selected according to the research design in this research study. Besides, the quantitative method was chosen and questioned as the primary source of the research study. The research location was conducted in Lampung, Indonesia and Ho Chi Minh City, Vietnam. 200 questionnaires will be the total population distributed electronically, namely google form by researchers to respondents. SPSS will be used as a statistical tool to analyze data.

