

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Instrument

In this project there are two instrument to support this project. The instrument used to support this project is:

1. Software

- a) The operating system that will used is Windows 10
- b) Notepad ++

2. Hardware

- a) The processor that will used is AMD A4-3330MX APU with Radeon(tm)
- b) 500 GB hard driver
- c) 8GB RAM

3.2 Data Collection

The training and testing data are obtained from form questionnaire through customer. The data will converted into excel format or csv before imported into database MYSQL. The data are taken from google doc and the total data is 400 data for the training data.

3.3 Research Procedure

In this project there are several stages in the research procedure. This is done so that the project made is more focused and organized. The following are several stages of the research procedure.

1.Literature study

Collecting 10 journals related to Data Mining, Decision Tree C45, K-Nearest Neighbor, Cross Validation and Confusion Matrix then make analysis how the problem could solved using those implementation

2. Data Collection

Collecting data from Customers using Google Docs questionnaire form. The target of this project is people who visited in coffee shop. Every visitor who comes to the coffee shop will be given a questionnaire related to the menu they ordered

3. System Design

The system design should contain comparison between two algorithm to find where is the best algorithm to make recommendation menu. The step design will be described with a flow chart. This process will help to keep how the procedure work is.



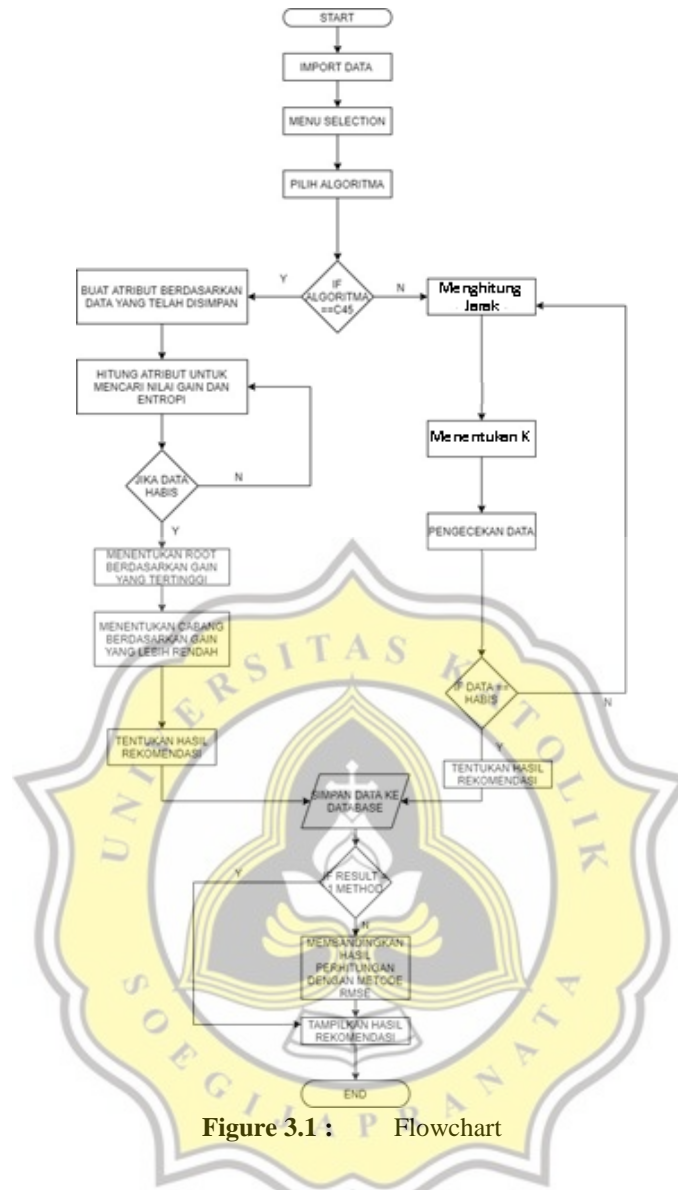


Figure 3.1 : Flowchart

Figure 3.1 the first step to do is import data to database. After it done, user must chose variable and algorithm and submit process. Preprocessing data will do automatically since user selected variable. The result from this project will automatically save to database. Each result of algorithm will be compared in the end classification process.

4.Coding and Testing

After data already obtained, first thing to do is convert it to csv or excel format. After that import the data into MYSQL database. Next step is the preprocessing stage which is contain filtering and sorting according to the variables will be calculated and algorithm used. After choose method and variable the data will be processing according to what was chosen. Then the program will be compared the result between two algorithm using confusion matrix method. This project program codes is using PHP programming language and laravel framework.

5.Report

Conclusion from this project is K-Nearest Neighbor is better than Decision Tree C45 to make classification. The comparison from two algorithm is using confusion matrix to get percentage of accuracy. K-Nearest Neighbor get accuracy higher than Decision Tree C45. K-Nearest Neighbor can get 100% accuracy while Decision Tree c45 just get around 80%.

