CHAPTER 3
RESEARCH METHODOLOGY

Of both journals has been taken then there is the idea that makes the authors are interested in combining both of these journals. The authors wished to classify sentences based on the value of hash.

3.1 Literature Study

The initial stage in the creation of this project was to read journals related to the project. A journal that I use in literature study I was using the winnowing algorithm Journal of Solar Priantara, Wayan dkk (2011) and k-means clustering journal from Asroni, Ronald Adrian (2015).

3.2 Problem Analysis

The second stage in the making of this project is to analyze the problem. The winnowing of the algorithm has been implemented there is a conclusion that the results of the project the level of similarity is too low. With a shortage of that I want to try to make the level of similarity based on grouping based on the value of its proximity.

3.3 Design

The third stage in the making design, this project is to create a design that fits. The system starts by inserting a sentence tekutul then remove punctuation and spacing as well as revamp into lowercase uppercase. Next change sentences into hash value. After the hash value is formed and then determine the centroid values randomly from the hash value that has already been formed. Hash value that has developed later than one by one with the centroid proximity value so formed using the distance between the centroid proximity value. Classify each hash value comparisons with the proximity corresponds to the number of groups to be formed.
3.4 Implementation and Testing

The fourth stage in the manufacture of the system. This stage is implementation and testing. The program will be created using geedit. The program will be shown in the form of testing phase in the GUI. Try going to do when the program is run the error is still present or not.