

CHAPTER 3

RESEARCH METHODOLOGY

In this research, there are several steps to be taken.

1. Studies Literature

Find references in journals or books about sentiment analysis, data mining and data mining classification.

2. Identify Problem

Find and analyze what problems occur in determining sentiment analysis on the questionnaire. In this project, the authors found the problem of how to determine the questionnaire that has a positive and negative response. After get a problem, the author looks for algorithms to solve the problem.

3. Collecting Data

Collect training data and test data before applying the algorithm. Data is taken from crawling data from BMSI Unika Soegijapranata and web which have been made before. Data is stored in txt format.

4. Prepare Data

Prepare data before processing sentiment analysis contained in the questionnaire. Data is processed by preprocessing method. There are 3 stages in preprocessing process that is tokenization, stopword removal and stemming. After completion of the preprocessing step, the data is ready to be classified with K Nearest Neighbors algorithm.

5. Classifying Data

Classification data using K Nearest Neighbor to get the classification result of questionnaire which has positive response and negative response.

6. Testing Data

Examine existing data with different amount of training data and k value at K Nearest Neighbors. To know the difference of result of each data process.

7. Report Writing

Write the conclusions about the classification using K Nearest Neighbors, as well as how much accuracy is obtained using K Nearest Neighbors to determine the questionnaire with positive and negative responses.

