BAB I

Natural disasters in Indonesia very familiar, examples of natural disasters that have occurred are tsunamis, earthquakes, landslides, and other natural events. According to Law Number 24 of 2007 concerning Disaster Management, natural disasters are events that occur due to natural events [1]. There are many ways to solve the problem of natural disasters, fast information can be a solution. Nowadays information spreads very easily on the internet through social media, social media especially X has become a social media that is widely used to share information. It is also important to analyze the hashtags given to each post in X so that the hashtags are relevant to the post, so it is also very important to do a classification of hashtags that are relevant and irrelevant to the disaster. Hashtag classification can help some parties such as the government or disaster management agencies to get more accurate information. From this information, direct monitoring, coordination and sending assistance needed by communities affected by natural disasters can be done.

In this research, a sentiment analysis of natural disasters will be carried out, especially in the natural disasters of the Cianjur earthquake, floods, and erupting Ruang mountains. And the data to be used are hastags from X media. Sentiment analysis or opinion mining is used one of them to support decision making by extracting and analyzing text, and identifying positive and negative opinions [2], This research will produce positive and negative classifications, the analysis will also be carried out using the Support Vector Machine (SVM) method and also using the Term Frequency-Inverse Document Frequency (TF-IDF) feature. Some sentiment analysis research uses Support Vector Machine, this method is the best among several other methods because it is able to compute data with high dimensions so that the resulting accuracy rate is better [3]. Data from X social media will also be categorized into food, water, medical emergency, shelter, electricity, evacuation routes, clothes, and neutral. With these categories, it will help the government and disaster management agencies to know what is needed by victims of natural disasters based on reports on social media X from the community. So that these parties can be more effective in providing assistance.

Machine learning research on sentiment analysis is general used in recent years, such as research [4] doing a comparison between Support Vector Machine and Logistic Regression for disaster classification. between Support Vector Machine and Logistic Regression for natural disaster classification. And the dataset used comes from twitter social media. After labelling and feature extraction of 1309 disaster tweets, 89 emergency tweets, 168 non-emergency tweets were obtained, and 1052 irrelevant tweets. and this study's results demonstrate that Support Vector Machine and Logistic Regression are used to classify natural disasters. The results of this study show that Support Vector Machine has an accuracy of 80.41%, while the Logistic Regression method has an accuracy of 63.36%,

In research [5], by combining feature extraction methods word2vec classification Support Vector Machine with a total data of 10,000 reviews data taken from movie reviews. the best result with 78.74% f1-score, the best result is due to the higher the higher the dimension, the results of the vocabulary formed by the word2vec model will be the higher the variation, thus causing the accuracy to optimizing.

In research [6], sentiment analysis regarding post-disaster tweets that have not been classified. Therefore, the classification of data into positive or negative categories using the Naïve Bayes Classifier algorithm. Bayes Classifier algorithm. From the results of the classification that has been carried out, the results are positive results can mean that assistance based on these categories has been fulfilled while negative is less or needed. This research also uses The results of this test obtained an accuracy value for unigram of 76.67%, 84.44%, 90.00% and 93.33%. While the accuracy value for bigram is 64.17%, 68.89%, 75.00%, 86.67%. From the four tests, the results obtained accuracy is for unigram 93.33% and bigram 86.67%. Consequently, it can be said that the accuracy value of unigram is higher than bigram.