

# CHAPTER 1

## INTRODUCTION

### 1.1 Background

With the development of technology today it makes it easier for us to exchange opinions or opinions through social media. This will certainly cause a problem, namely differences of opinion or opinion towards fellow social media users. Therefore this project was created to make it easier to analyze the opinions of social media users by using sentiment analysis. Sentiment Analysis itself is the process of analyzing various data in the form of opinions or views so that conclusions are generated from various existing opinions. The problem to be resolved in this project is how the opinion of Twitter social media users on the current social distancing. This problem occurs because of the Covid-19 virus which is currently endemic, one way to avoid or reduce victims of this virus, the government is implementing social distancing to reduce or prevent the spread of the Covid-19 virus.

The way to solve this problem is by using the Vector Space Model algorithm and Naive Bayes algorithm. Which is then implemented into Sentiment Analysis to get maximum results. The first step of data is taken by scraping from Twitter first. Then the data will be processed by the preprocessing method before being completed using the Vector Space Model algorithm which will produce positive or negative final results. After that, it will be evaluated by calculating accuracy, precision, recall and f-score.

The final result of this project is whether the opinion of Twitter users on this social distancing problem supports the government to impose social distancing or rejects social distancing. And **whether** the Vector Space Model algorithm or Naive Bayes algorithm has better performance in certain ways. The final results of this project can be used for further research if needed. Or it can be tested using other algorithms.

### 1.2 Problem Formulation

1. What is the opinion of twitter users on this issue
2. **Whether** this method can be used to solve this problem
3. How good is this algorithm by calculating accuracy,precision, recall and f1-score

### 1.3 Scope

1. Only discussing social distancing tweet, not about the virus
2. The final results issued is positive or negative
3. The data used are only from twitter

### 1.4 Objective

1. Knowing twitter user **opinion** on this issue
2. Whether this method can solve it
3. Knowing how good this algorithm is from calculating accuracy, precision, recall and f1-score

