Chapter I INTRODUCTION

I.1 Background

In the digital era, with the development of the times and information technology today brings very significant changes to each individual, the more rapid the dependence of individuals with social media becomes a lifestyle that is common every day. Social media has many impacts on various kinds of information with specific fields such as politics, economics, education, health, etc. easily accessed quickly, and the latest information sources are very open because of the very rapid development of information technology in the modern world era. The information is easily obtained through various platforms found on the internet, such as websites, applications, social media, one of the tools that are always used as a liaison in multiple things that have many benefits that exist. Social networking media like Twitter, Facebook, and Youtube are some of the most popular communication media tools in the community today (Aliandu, 2012; Kumar dan Sebastian, 2012).

Twitter is one of the fifth most social media in use in 2020 (Statista, 2016). Twitter is used to express their thoughts and opinions about one thing or another. Advances in technological civilization make it easier for individuals to search for information, giving news no longer need to use letters and even information from around the world can get to all corners of this country. Data from 2015 shows there are more than 285 million Twitter users with 400 tweets users every day until now. Twitter is an essential domain of information (Ilmiah, 2017). Twitter provides a public API to access data on Twitter using a programming language, such as Python, which has a Tweepy library to access the Twitter API (H. Manguri, N. Ramadhan, & R. Mohammed Amin, 2020). Indonesia ranks third in the number of Twitter users in the world, and third in Southeast Asia with 19.5 million Twitter users in Indonesia (Kominfo.go.id) Twitter users are increasing from year to year and are increasingly interested in the community, the increase in the use of Twitter can be seen in figure 1.1 :

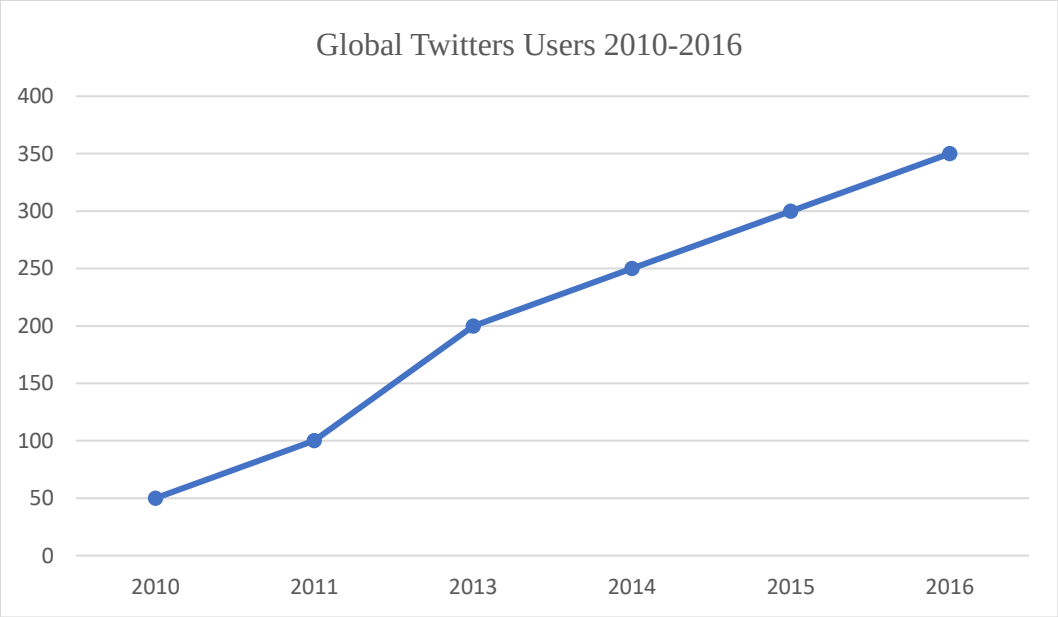

Figure I-1. Global Twitter Users 2010-2016 (Katadata)

According to Katadata, until in 2020, social media Twitter still enjoyed by the community from various groups, can be seen in the diagram above that displays Twitter.

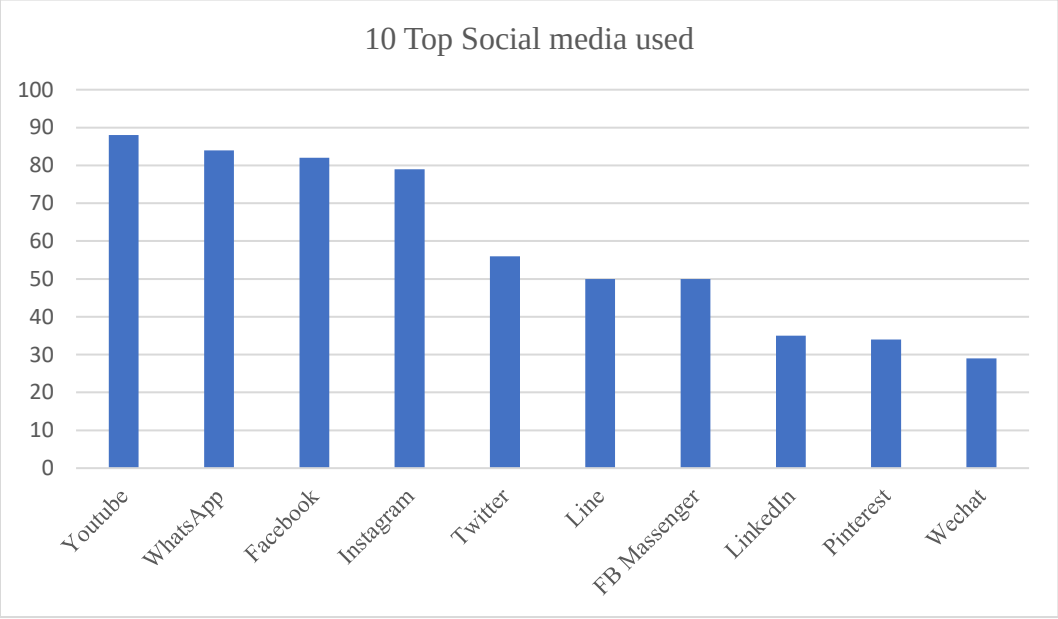

Figure I-2. Social Media That Is Often Used (Katadata)

In the early 2020s, the world was shocked by a pandemic called COVID-19 or also known as the Coronavirus, the World Health Organization (WHO) states that this Virus has become a global pandemic, COVID-19 is a severe respiratory disease caused by the Coronaviridae Virus with

symptoms of cough, fever, fatigue. This Virus is a derivative of (SARS-CoV-2) (Chamola et,al 2020). Indonesia is the fourth country estimated to be affected by COVID-19 in a long period (Djalante et al., 2020).

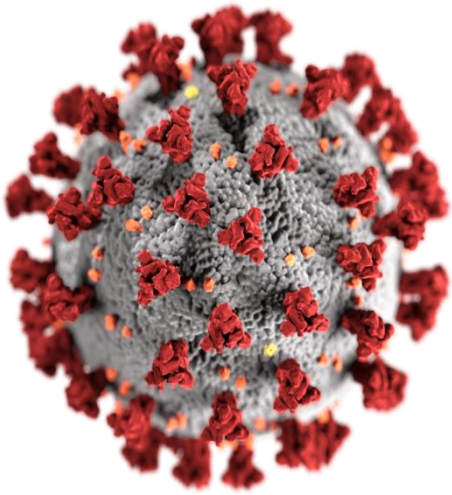

Figure I-3. COVID-19 (Wikipedia)

So far, there has not been much research on the analysis of special sentiment analysis about the spread of COVID-19 in Indonesia. This study takes data from Twitter with trending topics on COVID-19, the data used is data made using the Twitter API, data used as training and testing data. To find out how people reacted to COVID-19. One way to find out the community reaction to COVID-19 to conduct sentiment analysis can use data from social media one of them is Twitter. Sentiment analysis is opinion and logic expressed via textual (Zulfa & Winarko, 2017).

One way to do sentiment analysis can be done using data from social media. Thousands of submissions occur every day on every social media. Everyone can express their opinions through social media freely. In the research sentiment analysis Algorithm that is commonly used for the classification of sentiment analysis is Support Vector Machine, Decision Tree, but in this study using the Naïve Bayes Classifier because the Naïve Bayes is a simple model for text categories and produces precise accuracy (Sentiment & Work, 2009).

I.2 Problem Formulation

Based on the final project background, the problem to be investigated is as follows:

1. How to measure sentiment analysis with the Naïve Bayes algorithm with COVID-19 data?
2. What is the accuracy of the Naïve Bayes algorithm in sentiment analysis in the COVID-19 case?
3. How do the Indonesian people react to COVID-19 on social media based on sentiment analysis using the Naïve Bayes algorithm?

I.3 Research Objective

The formulation of the research in this Final Project are as follows:

1. Know how Twitter user sentiment analysis will be of the COVID-19 in Indonesia using the Naïve Bayes algorithm;
2. How to find out the level of accuracy of the Naïve Bayes algorithm in analyzing sentiments regarding the Covid-19 case;
3. Find out the impact caused by COVID-19 in Indonesia on Social media twitter in Indonesia by using the results of the Naïve Baye algorithm.

I.4 Research Benefit

The formulation of the research benefit in this Final Project are as follows:

1. For government

   Knowing the response of the Indonesian people on Twitter social media And as an input for the government in firmly handling the COVID-19 pandemic based on the results of sentiment analysis
2. For public

   Knowing the response of the Indonesian people on social media Twitter about the COVID-19
3. For research

   Carried out sentiment analysis based on COVID-19 comments on Twitter and Help provide references on sentiment analysis for further related research

I.5 Scope

To avoid material distortion and to broaden of the subject matter in this study, the limitations of the problem are:

1. In this study, the case to be taken is the reaction of the people in Indonesia related to COVID-19;
2. Using social media Twitter with top trending from March to May;
3. The algorithm used in this study is Naïve Bayes.