CHAPTER 4
ANALYSIS AND DESIGN

4.1. Analysis

This study uses the SVM algorithm to determine negative, neutral and positive sentiment in the classification of sentiment analysis, comments and opinions on the twitter application. In this study the programming language used is Python. By using the google colab text editor to run the project. The dataset is taken from opinions and comments that come from the twitter application. To get this data, you must have a twitter account and have registered as a developer account to get an API key. The data is taken using the crawling method. The data obtained are datetime, name, and tweet. The following is an example of the data set obtained after crawling the data :

Table 4.1. : dataset results obtained using the crawling method

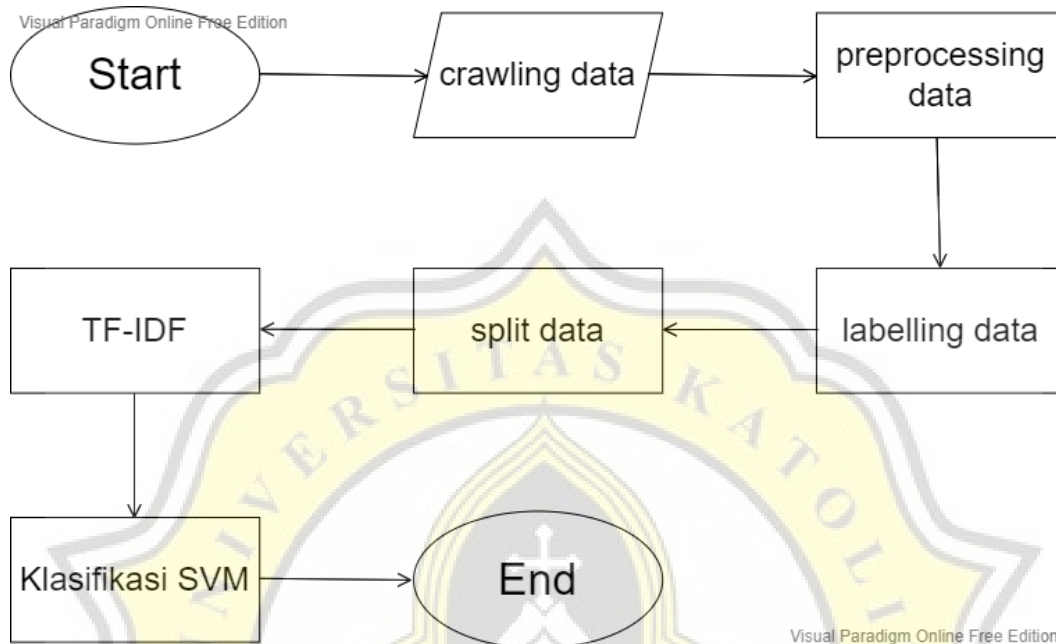| no	datetime	name	Tweet
1.	2021-11-23 13:06:35	polresjembrana	Bhabinkamtibmas Kelurahan Tegalcangkring, Polsek Mendoyo Pengamanan Dan Pemantauan Vaksin Covid 19 Door to Door. https://t.co/GtkalyHXlH
2.	2021-11-23 13:05:54	kemitris	@AldoBabeh Anies gagal dalam urusan vaksin, masih kurang mumpuni jadi pemimpin. Walau debatable, covid 19 bukan wab… https://t.co/78ToKcfr6r
3.	2021-11-23 13:05:37	tvOneNews	Mobil Vaksin Covid-19 Keliling, Ridwan Kamil: Kami Mengapresiasi Armada Vaksin Ini https://t.co/S1vxaAAtZa
4.	2021-11-23 13:05:29	VvVwWwLq	RT @TedHilbert: Mau tahu kenapa orang2 seperti @dr_koko28 @blogdokter @tonangardyanto @__Sridiana_3va @drpriono1 dan masih banyak lagi meny…
5.	2021-11-23 13:06:35	polresjembrana	Bhabinkamtibmas Kelurahan Tegalcangkring, Polsek Mendoyo Pengamanan Dan Pemantauan Vaksin Covid 19 Door to Door. https://t.co/GtkalyHXlH

4.2. Design

Figure 4.1 : flowchart of the process to be carried out on the project

The first step in this research is to take the dataset using data crawling, then the data that has been obtained must go through the pre-processing stage, namely the case folding stage, tokenization, filtering and stopwords. After that, it goes into the labelling process to separate negative, neutral and positive sentiment data. Then split the data to get test data and train data. Then, enter the TF-IDF process, and after that enter the SVM Algorithm process.