CHAPTER 1
INTRODUCTION

1.1. Background

Since it was officially announced on February 15, 2022 and its is located in part of Penajam Paser Utara Regency and partly in Kutai Kartanegara Regency. The mega project of the new capital city is already reaping the pros and cons. In the other side covid cases that exploded in Indonesia in the same year, the budget allocation has also become a controversy in the middle of the pandemic storm. People think that the government is wrongly targeted in budget allocations and are asked to focus more on dealing with Covid-19. On the twitter social media platform, this topic is quite popular and widely discussed. To see the public's response to this issue whether it tends to be positive, negative, or neutral, a sentiment classification will be carried out with the aim of providing an overview of the response of the general public to this government mega plan.

Based on previous exposure, this sentiment analysis was carried out on tweets that were crawled in the time span of June 5 – 11, 2022. Then the data is preprocessed to remove characters, links, and other things that can interfere with the weighting of sentiment. Then it will be labeled into 3 groups, namely positive, negative or neutral. Weighting is done using an English-language lexicon. Finally, it is analyzed with the Support Vector Machine algorithm..

Sentiment analysis on the twitter platform with the topic of the new capital city. With data in the form of tweets in Indonesian, the sentiment is determined by the lexicon method which will be classified using Support Vector Machine to see the effectiveness of the sentiment results. After that the classificatiion result tested again uses K-Fold Cross Validation

1.2. Problem Formulation

1. How effective Support Vector Machine (SVM) in classification of sentiment analysis on tweet data based on the lexicon method ?

2. How much data will be used for this research ?

1.3. Scope

1. The data used is limited to the IKN topic with the keywords #IKN, #ibukotabaru, IKN
2. The data is in Indonesian tweet
3. Data tweet is mining with twitter API using tweepy
4. Sentiment on the data is obtained by using the Indonesian lexicon
5. The classification method used is Support Vector Machine (SVM)

1.4. Objective

The purpose of this research is to prove that the Support Vector Machine method can be used to classify negative sentiments or positive sentiments in sentiment analysis of tweets based on the lexicon method about the capital city of a new country. This is the methodology used to find out the trend on the topic of this new capital city easily.