

Sentiment Analysis on Social Media Using Long Short-Term Memory and Word2Vec Feature Expansion Methods with Adam Optimization

Sanabila Khoirunnisa ¹, Erwin Budi Setiawan ^{2*}

erwinbudisetiawan@telkomuniversity.ac.id

^{1,2}School Of Computing, Telkom University
Bandung

Abstract Twitter is one of Indonesia's most popular social media, so it has many users. The intensity of Twitter use can be used to carry out sentiment analysis related to topics being widely discussed, especially regarding the 2024 Indonesian presidential election. To understand public views, public opinion is divided from text data into positive and negative polarities to measure public sentiment. The classification model uses Long Short-Term Memory (LSTM) for feature extraction, utilizing TF-IDF. In addition, this model also combines Word2Vec based on the Indonews corpus, which contains 142,545 articles for feature expansion. This model is further optimized using the Adam optimization technique to improve accuracy. By using a dataset of 37,391 data, the results of this research obtained an accuracy score of 83.04% and an f1 score of 82.62%. This is an increase in accuracy of 11.22%; for the f1 score, it was a 10.92% increase from the baseline. This indicates that the classification model using Long Short-Term Memory (LSTM) with the application of TF-IDF as feature extraction, Word2Vec as feature expansion, and Adam optimization successfully produced optimal sentiment predictions regarding the 2024 Indonesian Presidential Election.

Keywords: Sentiment Analysis, LSTM, TF-IDF, Word2Vec, Adam, Twitter

Article info: *submitted January 1, 2020, revised February 2, 2020, accepted March 15, 2020*

1. Introduction

In the current digital era, information technology is developing quite rapidly. With the development of information, society can communicate more quickly and efficiently by using social media. Social media is an online media that users can use easily to participate, share, and create content like social networks. Wrong One social networking site that is often used is Twitter. Based on statistical data from Statista [1], Indonesia is ranked fifth in the world, with the number of users reaching 24 million users. With so many Twitter users, this platform has become one of the fastest digital media disseminating information.

Twitter is a social media platform that people can use freely to express their opinions. People can use Twitter to send messages or what are usually called tweets. One of the topics that is being widely discussed is tweets related to the 2024 Indonesian presidential candidates. The existence of tweets causes people to have views on trending topics on Twitter. However, the more information a user receives, the more tweets there will be, which sometimes have ambiguous meanings, thus affecting understanding of the tweet. This has led to various proactive and

counter-responses from Twitter users. Emotions in textual form are called sentiment analysis[2].

Sentiment analysis or opinion mining is a branch of research text mining. Sentiment analysis can be used to analyze opinions, emotions, and written assessments on specific topics using processing techniques in natural language[3]. Sentiment analysis is related to the polarity identification sentiment of a text, which states whether the text is positive, neutral, or negative. Identification of hidden sentiments refers to emotions in between anger, happiness, and annoyance [4]. There are several branches of sentiment analysis, one of which is aspect-based sentiment analysis (ABSA). This analysis makes it possible to obtain more detailed information based on each aspect[5]. There are many methods for conducting sentiment analysis for a topic.

One of the sentiment analysis processing methods that can be used is deep learning. Deep learning allows computers to learn complex concepts by breaking them down into more straightforward concepts to gain a broader understanding[6]. In recent years, there have been significant advances in deep learning technology in the field of natural language processing, and many online text sentiment analysis methods based on deep learning have been proposed[7].