1. Introduction

1.1. Background

The development of information and communication technology certainly has a great influence in this current era, not only in the real world but in the form of cyberspace, especially on social media. It can easily and quickly disseminate information connected to the internet network (Togay et al., 2022). Social media is a software that aims to enable users to indoctrinate, communicate, argue, share, network and other activities (Sriyano & Setiawan, 2021). For example, on the social media of Facebook, Instagram, and Twitter, users can easily and quickly spread information that changes the topic of being discussed, such as discussing the community environment, politics up to the entertainment industry. As for (Sriyano & Setiawan, 2021) stated that the hoax news is currently developing in social media circles (Dirjen *et al.,* 2017).

Fake news can provide the wrong information for the users so that those who read the news can believe that this news is misleading and can scare the users who receive it (Batoebara *et al.,* 2020a) This news seeks to deceive and trap the public and adversely affect users who easily believe (Batoebara *et al.,* 2020b). Social media is an online medium, whose users can easily participate, share, argue, and create content in the form of blogs, social networks, wikis, forums, and virtual worlds (Istiani & Islamy, 2020). Twitter is one of the social media that is widely used by Indonesian people (Dirjen *et al.*, 2017).

Twitter is a *micro-blogging* social media founded by Jack Dorsey in March 2006 and started to be used on July 2006. Twitter has the uniqueness of only being able to post a text with a maximum of 140 characters namely a *tweet.* This media often also uses more than one language or so-called bilingual language (use two or more languages) [3][11]. On Twitter, users can disseminate information through tweets (Tineges et al., 2020). The more *tweets*, the wider the information obtained. However, some fake news on *tweets* often occurs. This will have a very bad impact on recipients who easily believe its information, for example on the impact of COVID-19 which stated the number of hoax news scaring the public and can attack people's psychology (Yunanto *et al.,* 2021).

Hoax news is compiled not based on facts (Parewe *et al.*, 2021a). In 2019, Ananthi *et al.* detected hoax news and real news using deep learning (Parewe *et al.*, 2021b). Moreover, in their research, Ananthi *et al.* compared machine learning and deep learning techniques. The dataset taken comes from Kaggle.com website of hoax news and real news datasets. The methods used *K-Nearest Neighbor, Decision Tree, Naïve Bayes, Random Forest,* CNN (*Convolutional Neural Network*), and LSTM (*Long Short Term Memory)* using *Term* Frequency *Inverse Document Frequency* (TF-IDF) weighting. The results of deep learning techniques through the CNN and LSTM methods have a better accuracy value than using mechine learning techniques (Kurniawan & Mustikasari, 2021).

Based on the description above, the impact caused by hoax news is very detrimental to many parties. Therefore related research on the detection of hoax news is mostly carried out using the *Decision Tree*, *Support Vector* (SVM), *Self-Organizing Map,* and *Convolutional Neural Network* methods. (CNN). In this study, in order to detect hoax news using the CNN method, it is actually more effective and efficient using the CNN method to classify documents, but the author tries to use CNN in text form to find out whether CNN can work well in managing text or not. This study used the weightings of Term Frequency-Inverse Document Frequency (TF-IDF) and Term Frequency-Relevance Frequency (TF-RF) weighting (Farid *et al.*, 2020a).

1.2. Topics and The Limitations

The topic discussed in this study is the fake news (*hoaxes*) detection of *tweets* on Twitter social media using the *Convolutional Neural Network* (CNN) method by comparing the use of weighting features *of Term Frequency Inverse Document* Frequency (TF-IDF) and *Term Frequency-Relevance Frequency* (TF-RF). The limitations on the study used datasets based on tweets on Twitter social media using the hashtag "covid-19 vaccine" and the dataset retrieved only using Indonesian language.

1.3. Objective

The purpose of this study is to acquire the optimal CNN architecture in identifying hoax news of tweets on Twitter social media and the comparison of the success rate results of the CNN classification system with the TF-IDF weighting feature for CNN classification system with TF-RF weighting in identifying hoax news on Twitter social media.

1.4. Writing Organization

The research section is structured as the section in chapter 1 which contains an introduction, the section in Chapter 2 contains an explanation of the related studies, the section in Chapter 3 contains an explanation of the methods used and their application, the section in Chapter 4 contains the results of its testing and analysis, and the section in Chapter 5 contains the conclusions of this study.