

1. Introduction

Personality plays a significant role in human's lives as social beings. Each personality shows the behaviors of each individual in socializing. According to Florance Littauer, in her book titled *Personality Plus for Parents*, she mentioned that there are four personality types, namely sanguine, choleric, apathetic, and melancholic [1]. Moreover, there is also a Big Five Personality describes that describes the human personalities that can be divided into five and called OCEAN. OCEAN is an abbreviation for Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism [2].

The importance of human personalities is shown by the prevalence of personality tests as an important requirement in various recruitment processes, ranging from organization recruitment, scholarship recruitment to job recruitment. According to tech.co, in 2014, an institution called Hyper Island released a survey of 500 CEOs, managing directors, hiring managers and employees in technology, communication and business [3] that 78% of the respondents believe that during recruitment processes, personality is deemed more important than skills [3].

But in reality, personality tests have several drawbacks. The excessive number of questions during a personality test can cause the results to be inaccurate, thus failing to represent the respondent's personality [4].

Therefore, it is necessary to have a method that can understand the personality of each respondent. There must be a platform to collect information that can detect personality to solve this problem. Social media like Twitter can be one of them. A research by Mohammad Zoqi Sarwani et al. [5] using the Naïve Bayes method shows that the tweets made by Twitter users represent their personalities.

In the classification process, machine learning-based approaches such as Naïve Bayes, Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) has been implemented. There are several weaknesses from these methods, Naïve Bayes does not consider the relationship between features, so it cannot take advantage of the relationship between speech tags, keywords and negation [6], while SVM classifies data into two classes. If there are more than two classes to be classified, modification is needed. If the training data used is very large, it can affect the training time and the required memory size [7]. Several researchers have developed a method that has a higher level of accuracy than machine learning techniques, which is deep learning. One of the methods of deep learning is Recurrent Neural Networks (RNN). RNN is a deep learning method with a small number of layers and has a backward connection flow [8]. Therefore, this RNN can be used for sequential data or sequence [6]. But the longer the input in a sequence, the more complex the RNN calculation will be. As a result, the capabilities of the RNN will be limited. The RNN architecture emerged to solve this problem or what is commonly known as vanishing gradient, the RNN architecture emerged, which is LSTM (Long Short-Term Memory). Another advantage of LSTM is that the RNN method has a very low error rate. It is proven by the research of Hassan and Mahmood [9] when classifying sentences using the IMDB and SSTb datasets. LSTM is superior to the seven other classification methods and has a low error rate [5].

Because the dataset is a combination of several tweets and has a connection with each other, LSTM is very suitable to be used in this research. This research also uses Word2Vec for dataset extraction.

Therefore, the difference between this system and the previous research is the chosen method that is expected to improve the performance of a model that can detect the personality of Twitter user tweets.

With this research, the personality detection process can be done easily and quickly. So the activities that require personality detection can be carried out easily and quickly.