

1. INTRODUCTION

Depression, that also goes by the name major depressive disorder, is commonly found in people with anxiety disorders. The common symptoms of depression include anxiety attacks, sleep disturbances, loss of appetite, ongoing sadness (1 month), and loss of interest in favorite activities [1], [2], [3]. Depression affects 280 million people, or 3.8% of the population, with a larger rate among women. In addition, 700,000 individuals die by suicide each year, making suicide the top cause of mortality among 15-19 year olds [4] and positioned 13th among the biggest causes of mortality globally [5]. There were over 300 million individuals worldwide who experienced In 2019, the World Health Organization (WHO) reported. Indonesia accounted for 15.6 million of these individuals. However, it seems that only about 9% of patients get professional therapy based on those statistics [6].

Concurrently, social media has evolved into a digital platform for communication and sharing [7]. Depression can be experienced by anyone, but often individuals and those closest to them do not realize it. Therefore, social media becomes a means of expressing their expressions and feelings [8]. Problematic social media users (PSMU) have a greater potential to develop depression compared to social media users (SMU) [9], which means that the high intensity of media use can lead to high levels of depression experienced [10], [11]. This issue increases the possibility that social media can serve as a supportive resource that mental health experts or practitioners can use when making diagnoses and looking for indicators of mental illnesses such as depression.

Interactions on X contain significant potential for identifying users' emotions and psychological conditions, including depression. Research [12] indicates that retweets amplify the dissemination of a content, hence enhancing the probability of it being read and engaged with by numerous users. The same holds true for information pertaining to psychological illnesses, since more interactions on tweets expressing feelings or thoughts related to depression may offer insights about the user's mental health. Quote tweets enable users to append personal remarks or thoughts to a post, potentially conveying their sentiments. By examining the frequency and characteristics of interaction patterns, signs of depression can be delineated. This research establishes a foundation for employing X interactions as data for identifying depressed states, as increased engagement with content related to mental health themes correlates with a higher probability of discerning an individual's feelings or depressive conditions.

The primary issue in conducting this research is the selection of an appropriate algorithm for identifying complex patterns. This research compares two machine learning techniques, Decision Tree and Random Forest, that have distinct qualities and benefits in classification tasks. The Decision Tree algorithm is, especially with large datasets [13]. Conversely, Random Forest, as an ensemble of multiple Decision Trees, improves accuracy and mitigates overfitting. Evaluating the ability of these two methods to detect signs of depression is critical to determining the most appropriate algorithm for mental health detection applications in social media [14].

Some research has been carried out, especially on platform X, to identify depression in its users. Research conducted by Marcello H. [13] using the CART model-based Decision Tree algorithm method with TF-IDF feature extraction with a maximum tree depth of 4 achieved an accuracy of 81.25% and an F1-score of 85.71%. In this study, changes in the amount of training data, TF-IDF feature values and tree depth values for the model significantly affects performance. so that changes must be made carefully so that the model does not overfit. In research conducted by Hatoon A. and Mourad Y. [15], the accuracy of Random Forest was 84.7% and Decision Tree was 77.5%. This research emphasizes how crucial feature selection is for identifying depression in machine learning algorithms. Based on previous studies, this indicates that the algorithms are able to provide better results in identifying depression through analyzing interaction patterns on social media. In research [16], the prediction of anxious depression was carried out using 100 user data obtained from scraping API X. This study used 3 classifiers, Random Forest obtained an accuracy value of 81.04. This indicates that Random Forest is able to identify anxious depression based on data on social media. Research conducted by Fiza A. et al. [17] to identify depression using Support Vector Machine, Random Forest, along with Bag of Words (BoW) feature extraction resulted in an accuracy of 0.73 and 0.77 respectively. In this study Random Forest excels because it has a better F1-score, the result is 0.78 for the class that is depressed and 0.76 for the class that is not. This indicates that Random Forest is able to classify the depression class accurately. In another research [18], depression detection was carried out from the Reddit platform using several machine learning algorithms. The Random Forest model excels and produces accuracy and F1-score of 0.877. This value was obtained after augmentation using SMOTE and Word2Vec feature extraction. Meanwhile, Decision Tree obtained the maximum accuracy achieved is 0.781, with an F1-score of 0.772.

Unlike previous research that focuses primarily on the content of user posts, this study examines the relationship between depression indicators and interaction patterns, such as the number of comments, likes, quotes, and retweeted quotes. Furthermore, this research not only contrasts performance, but also integrates three feature extractions with two distinct preprocessing scenarios, providing a more comprehensive approach to depression detection. It is anticipated that this research will assist mental health professionals and social media platforms in identifying early signs of depression in users. This research has the potential to serve as a foundation for creating more effective automated monitoring systems by evaluating the effects of machine learning algorithms and understanding interaction patterns as indicators of mental health. This will help prevent depression and increase awareness of the importance of social interaction in mental health.