1. Introduction

Social In this era of globalization, technology has advanced significantly, and the development of communication technology and the internet has made everyday life easier, more effective, and efficient. The internet has both positive and negative impacts on human life [1]. According to Aldiansyah Putra et al. [2], the number of social media users increased by 13.2% in January 2021, reaching 4.2 billion users compared to the same period the previous year.

Anxiety is a state experienced in response to threats that are either distal or uncertain, and involves changes in an individual's subjective state, behavior and physiology that facilitate detection of a potential threat within the environment [3]. A study by Mukhlasin shows that anxiety can be triggered by genetic factors, past experiences that affect psychological well-being, clinging too much on some substances, or too much time spent in social media. Lately, social media use is often linked to mental health problems [4]. Excessive use of social media has become a concern for researchers, parents, and society [4]. The American Psychological Association (APA) explains that anxiety among teenagers is specified by stress, constant worry, physical reaction such as increased heart rate, and concerns about uncertain things. Anxiety mainly appear from a fear of unknown objects or situations without a clear reason [5]. In this era of digitalization, social media has become a platform to exchange information and to acknowledge the condition of others that makes it one of the most relevant data source for analyzing a person's mental health [6]. X (formerly Twitter) is commonly used by society to express opinions on a range of issues, making it one of the most reliable data source for sentiment analysis to this day [6].

Two commonly used methods in sentiment classification are Naive Bayes (NB) and Support Vector Machine (SVM). Research performed by Ridho Fazal et al. [6] compares NB and SVM algorithms for sentiment analysis on X, indicated higher performance for SVM, with an accuracy of 88.52%, while NB achieved an accuracy of 82.51%. Another study by Widia Ningsih et al. [7] NB and SVM in sentiment analysis on X in the case of electric vehicle usage in Indonesia, resulting in an accuracy of 70.83% for SVM and 63.02% for NB. Another study by Moh. Aminullah Al Fachri and Ummi Athiyah [8] compared the NB and SVM algorithms to analyze the issue of expensive cooking oil using a dataset of 9,194, which showed an accuracy of 81.6% for both SVM and NB. In research by Rani Yunita et al. [9], SVM and NB algorithms were applied to sentiment analysis of ChatGPT on X, using a dataset of 1,000, with results showing 59% accuracy for SVM and 47% for NB. A study by Dewi Setiyawati and Nuri Cahyono [10] analyzed sentiments about smoking in Indonesia on X, using a dataset of 2,772, resulting in an accuracy of 60.8% for SVM and 62.1% for NB. However, these studies often encountered challenges with imbalanced datasets or inadequate feature extraction methods for this kind of task, leading to inconsistent accuracy and F1-scores. Therefore, this study is necessary to improve the performance of NB and SVM by addressing these limitations through the use of advanced feature extraction methods like TF-IDF and BoW and balancing techniques like Random Oversampler. Based on the above explanation, this study focuses on comparing the prediction algorithms for anxiety using NB and SVM to evaluate the performance of both. The aim of comparing these two methods is to determine which algorithm performs the best.

This study's contributions are twofold. First, the work presents a comparative analysis of the two popular algorithms in the context of mental health analytics. Then, it also points out the role of feature engineering in improving model performance and provides practical insight into future research and applications. Based on data from X, this study further develops anxiety detection and emphasizes the use of social media for early intervention in mental health.