

COMPARATIVE ANALYSIS OF MACHINE LEARNING MODELS FOR DEPRESSION DETECTION IN SOCIAL MEDIA TEXT

**M.U.K. Gunawardhna^{*}, K.B. Shalika, R.M.T.C. Bandara and
W.A.W.U. Wickramaarachchi**

*Department of Information Communication Technology, Rajarata University of Sri Lanka,
Mihintale, Sri Lanka.*

**udikagunawardhna@gmail.com*

Depression is a prevalent mental disorder that requires early detection. However, traditional clinical methods often face challenges in identifying their onset. Social media platforms, where users frequently share personal thoughts and emotions, provide a valuable source for detecting depressive symptoms. Although research in this area is expanding, relatively few studies have examined multiple machine learning (ML) models across different feature-extraction methods. Prior work has focused mainly on Twitter-based data and support vector machine (SVM) classifiers, with limited attention to other classifiers. Addressing this gap, this study evaluated three ML models, SVM, logistic regression (LR), and long short-term memory (LSTM) networks with dense layers on Facebook posts using two feature-extraction methods: term frequency-inverse document frequency (TF-IDF) and Global Vectors for Word Representation (GloVe) embeddings. A publicly available dataset was employed, and user demographics, such as age, age category and gender, were considered. Ground-truth depression labels were derived from an existing labelled dataset, where individuals diagnosed with depression were assigned a label of 1 and those without depression were assigned a label of 0. Eighty percent of the dataset was allocated to training and 20.0% to testing, ensuring class balance between depressed and non-depressed samples. Results indicated that TF-IDF with SVM achieved the highest accuracy of 95.5%, outperforming both LR and LSTM models, which each achieved an accuracy of 84.0%. This study demonstrates the effectiveness of combining SVM with TF-IDF in detecting depression in Facebook text and highlights the potential of extending research beyond Twitter-based studies. The findings contribute to the literature by systematically analysing multiple models and feature-extraction techniques using previously unexplored Facebook data, thereby supporting more robust and generalizable mental health monitoring through social media.

Keywords: Depression detection, Machine learning, Social media analysis, Support vector machine