



A Hybrid Method for Analysis of Sentiments and Opinions of Social Media users using Tweets dataset

Rubaina Nazeer^{1*}, Muhammad Sajid Maqbool¹

¹Department of Information Sciences, University of Education, Lahore

*Corresponding Author. Email: nrubaina@gmail.com

Received: 25 February 2022; Accepted: 3 March 2022

AID: 001-01-000005

Abstract: The COVID-19 epidemic swept the globe in 2019, having a significant influence on health, education, and the economy. New mutations like the Beta, Delta, and Omicron variations emerged as the coronavirus outbreak spread, frightening and alarmed the population. Around 6 million people have already perished from COVID-19 and its variations, according to World Meter. The "SARS-CoV-2" omicron strain was discovered on November 24, 2021, in South Africa, and has since spread to more than 57 nations. This study offers an analysis of people's attitudes and actions regarding the omicron variation. We suggest a technique for conducting sentiment analysis on information from Twitter about the omicron strain.

This study focused on processing omicron-related tweets using Python's NLP tools, and a dataset curated and optimized for feature extraction. This dataset served as the foundation for training machine learning models designed to categorize user emotional behavior into three distinct categories: Neutral, Negative, and Positive. The study employed six diverse machine learning classifiers - Naive Bayes, Random Forest, Decision Tree, Support Vector Machine, as well as one hybrid technique of Voting and Stacking classifiers by combining the classifiers. The central objective of the study is to accurately evaluate the predictive capabilities of these classifiers. Notably, the hybrid voting classifier exhibited a commendable performance accuracy of 85.33%, effectively categorizing the emotional behavior of users based on the omicron-related tweets. The ensemble stacking classifier, however, outperformed the other models with an even higher performance accuracy of 87.5%. These findings highlight the potential of ensemble techniques in improving the precision of sentiment analysis and emotional behavior classification, particularly in the context of real-time social media data processing related to current events like the omicron variant.

Keywords: Machine Learning; Sentiment Analysis; Opinion Mining; Ensemble Model; Covid-19.