

**SENTIMENT ANALYSIS ON FACEBOOK TOPIC USING NAIVE BAYES
ALGORITHM: A MACHINE LEARNING TECHNIQUE**

**A Thesis Presented to the Faculty of the
College of Communication and Information Technology
President Ramon Magsaysay State University
Castillejos, Zambales**

**In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science**

By:

**Tricia Erika A. Esguerra
Jasper A. Sadora
Martwin E. Yul**

June 2023

Republic of the Philippines
PRESIDENT RAMON MAGSAYSAY STATE UNIVERSITY
(Formerly Ramon Magsaysay Technological University)
 Castillejos Campus
 Castillejos, Zambales



APPROVAL SHEET

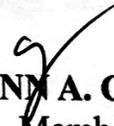
The thesis project entitled “**Sentiment Analysis on Facebook Topic Using Naïve Bayes Algorithm: A Machine Learning Technique**” was prepared and submitted by **Tricia Erika A. Esguerra, Jasper A. Sadora, and Martwin E. Yul** in partial fulfillment of the course requirements for the degree of **Bachelor of Science in Computer Science** has been examined and recommended for the oral examination.


MICHAEL G. ALBINO, MIT
 Thesis Adviser

Approved by the Panel of Examiners
 with a rating of 97 %


MARIE CELIA R. AGLIBOT, MSCS
 Chair


MICHAEL N. FARIN, MSCS
 Member


IRATUS GLENN A. CRUZ, LPT, MSCS
 Member

Accepted and approved as a requirement for the degree of
BACHELOR OF SCIENCE IN COMPUTER SCIENCE.

June 2023
 Date


IVY H. CASUPANAN, EdD
 Campus Director

ABSTRACT

Sentiment analysis has been useful in social media monitoring to automatically characterize the overall feeling or mood of consumers as reflected in social media toward a specific brand or company. It can also be used to determine the sentiments of people on specific topics through posts or tweets on their social media accounts.

The researchers aimed to gather Facebook statuses from various users' through Kaggle datasets and will apply machine learning techniques to extract the sentiments of users on the topic 'Artificial Intelligence'. Qualitative text analysis and applying Natural Language Processing Techniques (NLPT) were used to gather and automate the cleaning and data visualization of the sentiments.

Based on the result of the testing, the trained datasets revealed positive results that the SVM algorithm had 1.000 precision and recall and Radom Forest with the same value indicated that the two (2) algorithms gained a perfect score on the testing while the Naïve Bayes algorithm had a 0.940 recall and 0.840 precision result indicated six (6) errors on the trained model evaluation.

The researchers, therefore, concluded that the evaluation of the Naïve Bayes algorithms on the trained model had a minimal error as compared to the two different algorithms. The implementation of the sentiment analysis and trained datasets helped the researchers in determining the best algorithm for the model. It is recommended that future researchers should focus on other machine learning algorithms for the testing of precision and recall.

Keywords: *Datasets; Machine Learning Algorithms; Natural Language Processing;*

Visualization