



PROJECT REPORT
EVALUATION OF SENTIMENT ANALYSIS ON
SPOTIFY APPLICATION REVIEWS FROM THE PLAY
STORE USING SUPPORT VECTOR MACHINE AND
RANDOM FOREST ALGORITHMS

HENDRI PRATAMA
22.K1.0019

Faculty of Computer Science
Soegijapranata Catholic University

2026

ABSTRACT

Spotify, one of the world's largest music streaming sites, has experienced a significant surge in user feedback. This has resulted in a significant amount of collected review data, making manual analysis inefficient. Therefore, an automated method that can categorize reviews based on sentiment is needed. Therefore, The aim of this study is to assess how effective two algorithms, namely Support Vector Machine and Random Forest, are in analyzing sentiment from Spotify app user reviews sourced from the Play Store. This review data was obtained through data scraping techniques and totaled 234,589 reviews, which then entered a data cleaning process, which included cleaning, capitalization, tokenization, stop word removal, and lemmatization. Furthermore, at this stage, sentiment labeling was also carried out automatically using the VADER method, which divides the levels into three groups: positive, neutral, and negative. The data was then converted into numbers using the TF-IDF method to assign word weights before implementing the Support Vector Machine and Random Forest models. The final stage in evaluating model performance is measured through a confusion matrix that includes metrics such as accuracy, precision, recall, and f1 score. The results show that under imbalanced data conditions, the Support Vector Machine algorithm achieved the highest accuracy of 91.06%, slightly outperforming Random Forest, which obtained an accuracy of 88.18%. After applying the SMOTE method, the accuracy of both algorithms decreased, with Support Vector Machine reaching 88.86% and Random Forest achieving 88.46%.

Keywords: Sentiment Analysis, Spotify, Support Vector Machine, Random Forest