Sentiment Analysis of Online Learning Social Media Applications Using the K-Nearest Neighbors Method

Didit Rahmat Hartadi, S.Kom, M. T

Khairatuz Zahro

Study Program of Informatics Engineering

Major Of Information Technology

ABSTRACT

This study discusses sentiment analysis of user comments on the Pahamify application using the K-Nearest Neighbors (KNN) algorithm. A total of 1,200 Indonesian-language comments were collected, and after preprocessing, 1,173 labeled data (positive, negative, neutral) from expert annotation were obtained. The data was split into 80% for training and 20% for testing. The KNN model was trained to recognize sentiment patterns based on feature weighting using TF-IDF. To address class imbalance, the SMOTE method was applied, and Chi-Square testing was used for important feature selection. The model was tested using K values ranging from 1 to 18. The best result was achieved at K = 7, with an accuracy of 73% and the highest F1-Score of 80% in the positive class. These results indicate that the KNN model can effectively learn sentiment characteristics, particularly in recognizing positive sentiment.

Keywords: Sentiment Analysis, KNN, TF-IDF, SMOTE, Chi-Square, Pahamify