

Abstract

The increasing number of user reviews on the Google Play Store is a challenge in understanding user opinions and experiences with apps. One of the most discussed apps is Mobile Legends: Bang Bang (MLBB), a popular game with millions of downloads and reviews from Indonesian users. The problem faced is the limitation of conventional sentiment analysis models in understanding sentences and context simultaneously, making it less than optimal in analyzing user sentiment. This study proposed a comprehensive sentiment analysis system for MLBB application reviews, utilizing a hybrid CNN-LSTM architecture with a systematic optimization approach. A dataset comprising 30,000 balanced Indonesian user reviews was extracted from the Google Play Store using web scraping techniques and then processed through an extensive pre-processing pipeline, which included data cleaning, case folding, stopword removal, and stemming. Five experimental scenarios were conducted to optimize model performance through feature engineering and algorithmic enhancement. The baseline CNN-LSTM model achieved 71.97% accuracy, which was progressively improved through TF-IDF vectorization with optimal N-gram (1,2) configuration, max features optimization reaching 10,000 features, FastText embedding feature expansion using a 300-dimensional Indonesian pre-trained model, and optimizer selection experiments across five algorithms. The final optimized hybrid CNN-LSTM model, using the RMSprop, demonstrated a breakthrough performance of 88.84% accuracy with remarkable consistency (standard deviation of 0.000754), representing a 23.4% improvement over the baseline. This research contributes to the field of sentiment analysis, especially for game applications, by proving that a combined approach can produce a more accurate and reliable system for understanding user opinions.

Keywords: Classification, FastText, Hybrid CNN-LSTM, Sentiment Analysis, TF-IDF