Project Report

SENTIMENT ANALYSIS REVIEW COMMENT USING TRANSFORMER MODEL

VALENTINO PUTRA BUDI KUSUMO
21.K1.0005

Faculty of Computer Science
Soegijapranata Catholic University
2024

ABSTRACT

This research investigates the application of transformer-based models for sentiment analysis of movie reviews, emphasizing the role of preprocessing techniques. Traditional machine learning approaches have shown limited capability in capturing the emotional complexity of textual data. The study explores preprocessing methods such as case folding, symbol cleaning, tokenization, stopword removal, and stemming to evaluate their impact on the performance of the T5-small model integrated with a Multilayer perceptron (MLP). Using kaggle-sourced dataset, the research identifies as the optimal preprocessing combinations to maximize accuracy, precision, recall and F1-score. Experiments demonstrate that preprocessing enhances model;s generalization, reducing overfitting and improving performance metrics compared to unprocessed data. Grid search optimization identifies the best hyperparameters, achieving a peak validation accuracy of 70.56% with balanced precision and recall. The results. The results underline the necessity of tailored preprocessing for effective sentiment analysis while highlighting gaps for future research, such as dataset diversity and advanced hyperparameter tuning.

Keyword: Sentiment analysis, Transformer model, preprocessing, movie Reviews, Natural language processing