



ABSTRACT

Hybrid Sentiment Analysis Methods on Depression Detection

by

Marcellino Roland Putra Tamtama 19/444310/PA/19372

Depression has become an increasingly pressing global mental health concern, with many individuals expressing their emotional states through social media. Sentiment Analysis (SA), a natural language processing technique used to determine the emotional polarity of text, offers a valuable tool for identifying signs of depression based on user-generated content. While SA has traditionally been applied in domains such as marketing and product feedback, its application in mental health monitoring has shown significant promise. This research investigates a hybrid sentiment analysis framework for depression detection, combining machine learning algorithms with lexicon-based sentiment scoring to improve classification accuracy.

Specifically, the study employs Naive Bayes, and Random Forest as baseline machine learning models, which are further enhanced using sentiment lexicons created manually. A hybrid approach is developed through Lexicon-Enhanced Naive Bayes and Lexicon-Enhanced Random Forest, with an ensemble strategy based on weighted voting to synthesize model predictions. The models are evaluated on a publicly available depression-labeled dataset, with performance assessed using accuracy, precision, recall, and F1-score metrics.

By comparing traditional and hybrid methods, this study aims to identify the most effective model for detecting depressive content in social media posts. The findings are expected to contribute to the development of scalable, interpretable, and accurate systems for early-stage depression detection and digital mental health support.

Keywords : Sentiment Analysis, lexicon based approach, Naive Bayes, Random Forest, Ensemble