

## INTISARI

### PERBANDINGAN PERFORMANSI BILSTM, INDOBERT, SVM, DAN NAÏVE BAYES UNTUK KLASIFIKASI EMOSI PADA TEKS SINGKAT (TWEET) BERBAHASA INDONESIA

Oleh

Rheznandya Ardiyan Rahadika  
20/459277/PA/19938

Klasifikasi emosi pada teks singkat berbahasa Indonesia menjadi krusial untuk meningkatkan interaksi manusia-mesin, namun dihadapkan pada tantangan signifikan akibat karakteristik linguistik yang unik, seperti variasi bahasa informal dan keterbatasan dataset publik. Hal ini menimbulkan masalah dalam menentukan pendekatan model yang paling efektif dan efisien.

Untuk menjawab masalah tersebut, penelitian ini mengusulkan solusi berupa perbandingan performa model *deep learning* (BiLSTM dan IndoBERT) dengan model *machine learning* klasik (Support Vector Machine/SVM dan Naïve Bayes). Dengan menggunakan dataset tweet, dilakukan evaluasi komprehensif terhadap metrik performa, efisiensi komputasi, dan kemampuan analisis kontekstual untuk menemukan pendekatan yang paling optimal.

Hasil penelitian menunjukkan temuan yang kontras dengan ekspektasi umum: model klasik SVM secara konsisten menunjukkan performa tertinggi (Akurasi: 62,88%), mengungguli BiLSTM yang berperforma moderat (Akurasi: 58,68%). Sementara itu, IndoBERT, model yang paling kompleks, justru mengalami *model collapse* dan menjadi yang terendah dengan akurasi hanya 23,16%. Simpulan utama dari penelitian ini adalah sebuah "studi kasus peringatan" (*cautionary tale*) yang menegaskan bahwa untuk dataset informal berukuran terbatas, model klasik yang lebih sederhana dapat menjadi solusi yang lebih efektif dan efisien dibandingkan arsitektur *deep learning* yang kompleks tanpa adaptasi domain yang tepat.

**Kata kunci:** klasifikasi emosi, NLP, bahasa Indonesia, IndoBERT, BiLSTM, SVM, Naïve Bayes, analisis teks, pembelajaran mesin

## ABSTRACT

### PERFORMANCE COMPARISON OF BILSTM, INDOBERT, SVM, AND NAÏVE BAYES FOR EMOTION CLASSIFICATION IN SHORT TEXTS (TWEETS) IN INDONESIAN LANGUAGE

By

Rheznandya Ardiyan Rahadika  
20/459277/PA/19938

Emotion classification in informal Indonesian texts faces significant challenges due to unique linguistic characteristics and data limitations. This condition poses a problem in determining the most effective and efficient model approach for real-world use cases.

To address this problem, this study proposes a solution by comparing the performance of deep learning models (BiLSTM and IndoBERT) with classic machine learning models (Support Vector Machine/SVM and Naïve Bayes). Using a tweet dataset, a comprehensive evaluation was conducted on performance metrics, computational efficiency, and contextual analysis capabilities to find the most optimal approach.

The research results show findings that contrast with general expectations: the classic SVM model consistently demonstrated the highest performance (Accuracy: 62.88%), outperforming the moderately performing BiLSTM (Accuracy: 58.68%). Meanwhile, IndoBERT, the most complex model, experienced model collapse and ranked the lowest with an accuracy of only 23.16%. The main conclusion of this study is a "cautionary tale" which confirms that for informal datasets of limited size, a simpler classic model can be a more effective and efficient solution compared to a complex deep learning architecture without proper domain adaptation.

**Keywords:** emotion classification, NLP, Indonesian language, IndoBERT, BiLSTM, SVM, Naïve Bayes, text analysis, machine learning