ABSTRAK

PENERAPAN INDOBERT UNTUK ANALISIS SENTIMEN TAGAR #KABURAJADULU DI MEDIA SOSIAL X

Media sosial telah menjadi ruang publik yang dinamis untuk menyampaikan opini, termasuk dalam isu-isu sosial yang viral. Salah satunya adalah tagar #KaburAjaDulu yang ramai diperbincangkan dan menimbulkan beragam tanggapan masyarakat. Penelitian ini dilakukan karena belum banyak kajian akademis yang menelaah sentimen publik terhadap fenomena tersebut secara komputasional. Tujuan penelitian ini adalah menganalisis kecenderungan opini masyarakat serta mengevaluasi kinerja IndoBERT sebagai model Transformer dalam klasifikasi sentimen bahasa Indonesia.Penelitian menggunakan metode CRISP-DM yang meliputi pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran. Data dikumpulkan dari platform X (Twitter) melalui *web crawling*, kemudian diproses melalui pembersihan teks, normalisasi, penghapusan *stopword*, serta pelabelan otomatis dengan IndoBERTweet Sentiment. Untuk mengatasi ketidakseimbangan data, dilakukan augmentasi dengan *Contextual Word Embedding Augmentation*. Eksperimen dilakukan dengan dua skenario pembagian data (80:10:10 dan 70:20:10) serta variasi *epoch* dan *batch size*. Hasil penelitian menunjukkan model terbaik diperoleh pada skenario 80:10:10 dengan 10 *epoch* dan *batch size* 32, menghasilkan *precision* sebesar 93,03%, *recall* 93,02%, dan *F1-score* 93,03%. Model ini kemudian digunakan untuk mengklasifikasikan 8.000 data tambahan, yang menghasilkan 82,80% sentimen negatif dan 17,20% sentimen positif. Temuan ini memberikan gambaran bahwa opini publik terhadap tagar #KaburAjaDulu cenderung bernuansa negatif, yang dapat ditafsirkan sebagai bentuk kritik masyarakat terhadap isu tersebut. Secara metodologis, hasil ini juga membuktikan bahwa penyesuaian *hyperparameter* berperan penting dalam meningkatkan kinerja IndoBERT pada analisis sentimen isu sosial di media sosial.

Kata Kunci: Analisis Sentimen, IndoBERT, CRISP-DM, *Fine-tuning*, H*yperparameter Tuning*, Media Sosial

ABSTRACT

APPLICATION OF INDOBERT FOR SENTIMENT ANALYSIS OF THE #KABURAJADULU HASHTAG ON SOCIAL MEDIA X

Social media has become a dynamic public space for expressing opinions, including on viral social issues. One such issue is the hashtag #KaburAjaDulu, which has been widely discussed and has elicited a variety of responses from the public. This study was conducted because there have been few academic studies that have examined public sentiment toward this phenomenon using computational methods. The objective of this research is to analyze public opinion trends and evaluate the performance of IndoBERT as a Transformer model in Indonesian sentiment classification. The study employs the CRISP-DM methodology, which includes data understanding, data preparation, modeling, evaluation, and dissemination. Data was collected from the X platform (Twitter) through web crawling, then processed through text cleaning, normalization, stopword removal, and automatic labeling with IndoBERTweet Sentiment. To address data imbalance, augmentation was performed using Contextual Word Embedding Augmentation. The experiment was conducted with two data splitting scenarios (80:10:10 and 70:20:10) and variations in epoch and batch size. The research results indicate that the best model was obtained in the 80:10:10 scenario with 10 epochs and a batch size of 32, yielding a precision of 93.03%, recall of 93.02%, and an F1-score of 93.03%. This model was then used to classify 8,000 additional data points, resulting in 82.80% negative sentiment and 17.20% positive sentiment. These findings suggest that public opinion toward the hashtag #KaburAjaDulu tends to be negative, which can be interpreted as a form of societal criticism toward the issue. Methodologically, these results also demonstrate that hyperparameter tuning plays a crucial role in enhancing IndoBERT's performance in analyzing social issues on social media.

Keywords**:** *Sentiment Analysis, IndoBERT, CRISP-DM, Fine-tuning, Hyperparameter Tuning, Social Media*