

ABSTRAK

KLASIFIKASI SENTIMEN BERITA UNTUK PEMERINTAHAN DAERAH MENGUNAKAN METODE FINE-TUNING MODEL SENTENCE-BERT

Muhammad Zahidy Syawal – 1207050084

Jurusan Teknik Informatika

Penelitian ini bertujuan untuk membangun model klasifikasi sentimen terhadap berita pemerintahan daerah menggunakan algoritma Sentence-BERT (S-BERT). Studi dilakukan pada kumpulan berita tentang Kabupaten Purwakarta yang diperoleh dari dokumentasi internal Dinas Komunikasi dan Informatika (Diskominfo) Kabupaten Purwakarta. Total data yang digunakan sebanyak 1803 berita dan hasil Scraping GNews sebanyak 1879 Berita, kemudian diseleksi menjadi 894 entri dengan dua label sentimen utama, yaitu positif dan negatif.

Metode penelitian menggunakan pendekatan SEMMA (Sample, Explore, Modify, Model, Assess). Tahap preprocessing mencakup tokenisasi, penghapusan stopword, dan stemming menggunakan algoritma Nazief & Adriani. Evaluasi model dilakukan dengan metrik akurasi, precision, recall, dan F1-score. Proses representasi kalimat dilakukan menggunakan model *pre-trained* all-MiniLM-L6-v2 dan teknik SMOTE untuk menangani data tidak seimbang. Model klasifikasi dibangun menggunakan arsitektur *Deep Learning* dengan strategi *fine-tuning*. Hasil evaluasi akhir menunjukkan akurasi sebesar 87%. Analisis lebih lanjut menunjukkan kinerja yang kuat pada kelas positif (F1-score 0.93) namun lebih moderat pada kelas negatif, menghasilkan F1-score (macro average) sebesar 0.73. Hasil ini memperlihatkan bahwa model S-BERT efektif secara umum, namun kinerjanya masih dipengaruhi oleh karakteristik data yang tidak seimbang.

Analisis lebih lanjut menunjukkan bahwa sebagian besar kesalahan prediksi terjadi pada berita yang memiliki sentimen campuran atau konteks yang ambigu. Perbandingan dengan beberapa penelitian sebelumnya menunjukkan bahwa pendekatan ini memberikan performa yang kompetitif. Hasil penelitian ini memperlihatkan bahwa *S-BERT* merupakan metode yang efektif untuk klasifikasi sentimen berita pemerintahan daerah di Indonesia.

Kata kunci: analisis sentimen, SBERT, berita pemerintahan, Purwakarta, klasifikasi teks, SEMMA



ABSTRACT

A FINE-TUNING APPROACH USING SENTENCE-BERT ON SENTIMENT CLASSIFICATION FOR LOCAL GOVERNMENT NEWS

Muhammad Zahidy Syawal – 1207050084

Informatics Engineering

This study aims to develop a sentiment classification model for local government news using the Sentence-BERT (S-BERT) algorithm. The study was conducted on a collection of news articles about Purwakarta Regency, obtained from the internal documentation of the Department of Communication and Informatics (Diskominfo) of Purwakarta Regency. A total of 1,803 news articles and 1,879 news articles from GNews scraping were collected, which were then filtered to 894 entries with two main sentiment labels: positive and negative.

The research method employed the SEMMA approach (Sample, Explore, Modify, Model, Assess). The preprocessing stage included tokenization, stopwords removal, and stemming using the Nazief & Adriani algorithm. Model evaluation was carried out using accuracy, precision, recall, and F1-score metrics. Sentence representation was performed using the pre-trained model all-MiniLM-L6-v2, and the SMOTE technique was applied to handle imbalanced data. The classification model was built using a Deep Learning architecture with a fine-tuning strategy. The final evaluation results showed an accuracy of 87%. Further analysis indicated strong performance in the positive class (F1-score 0.93) but more moderate performance in the negative class, resulting in a macro average F1-score of 0.73. These results demonstrate that the S-BERT model is generally effective, although its performance is still influenced by the characteristics of imbalanced data.

Further analysis revealed that most misclassifications occurred in news articles with mixed sentiments or ambiguous contexts. A comparison with several previous studies showed that this approach delivers competitive performance. The findings indicate that S-BERT is an effective method for sentiment classification of local government news in Indonesia.

Keywords: *sentiment analysis, SBERT, government news, Purwakarta, text classification, SEMMA*