

ABSTRAK

Permasalahan utama dalam penelitian ini adalah bagaimana meningkatkan performa klasifikasi sentimen terhadap ulasan pelanggan restoran, khususnya dalam konteks bahasa Indonesia. Ulasan pelanggan mengandung ekspresi opini yang kompleks dan sering kali tidak eksplisit, sehingga memerlukan pendekatan pemrosesan bahasa alami yang mampu menangkap nuansa sentimen secara akurat. Model *IndoBERT* sebagai representasi *transformer-based language model* berbahasa Indonesia telah menunjukkan performa yang baik dalam berbagai tugas pemahaman bahasa, namun belum dioptimalkan secara spesifik untuk domain ulasan restoran. Oleh karena itu, penelitian ini berfokus pada pengaruh penggunaan fitur *N-Gram* dan optimasi *hyperparameter* dalam proses *fine-tuning* model *IndoBERT* untuk meningkatkan akurasi klasifikasi sentimen tiga kelas pada data ulasan restoran.

Metode penelitian ini melibatkan beberapa tahapan utama, dimulai dari pengumpulan data ulasan pelanggan dari Restoran Kampung Kecil Semarang. Data kemudian dilabeli secara otomatis menggunakan *lexicon* INSET untuk klasifikasi sentimen positif, netral, dan negatif. Dua pendekatan pelabelan diterapkan, yaitu dengan integrasi *N-Gram* dan tanpa *N-Gram*. Model *IndoBERT base-pl* digunakan sebagai fondasi, yang kemudian melalui proses *fine-tuning* dengan skema pencarian *hyperparameter* menggunakan metode *Grid Search*. Ruang pencarian dirancang berdasarkan rekomendasi dari Devlin et al., dengan mempertimbangkan kombinasi nilai *batch size*, *learning rate*, dan jumlah *epoch*. Evaluasi performa model dilakukan menggunakan metrik akurasi dan *F1-Score*.

Hasil penelitian menunjukkan bahwa penggunaan *N-Gram* memberikan dampak positif yang konsisten terhadap peningkatan performa model, baik pada tahap *pre-trained* maupun setelah *fine-tuning*. Model dengan *N-Gram* yang telah di-*fine-tune* mencapai akurasi sebesar 0,924429 dan *F1-Score* sebesar 0,924816, sedangkan model tanpa *N-Gram* hanya mencapai akurasi 0,846834 dan *F1-Score* 0,846685. Peningkatan performa ini menunjukkan bahwa *N-Gram* efektif dalam menangkap konteks frasa dan polaritas gabungan yang penting untuk tugas klasifikasi sentimen. Selain itu, proses *fine-tuning* dengan optimasi *hyperparameter* terbukti meningkatkan kinerja model secara signifikan dibandingkan model *pre-trained*. Penelitian ini memberikan kontribusi terhadap pengembangan sistem analisis sentimen yang lebih adaptif terhadap domain spesifik, serta mempertegas pentingnya pemanfaatan fitur linguistik tambahan dalam meningkatkan akurasi model berbasis *transformer*.

Kata Kunci: klasifikasi sentimen, IndoBERT, *N-Gram*, *fine-tuning*, optimasi *hyperparameter*

ABSTRACT

The main problem addressed in this study is how to improve sentiment classification performance on customer reviews in the context of the Indonesian language. Customer reviews often contain complex and implicit expressions of opinion, requiring a natural language processing approach that can accurately capture sentiment nuances. IndoBERT, as a transformer-based language model for Indonesian, has demonstrated promising performance across various language understanding tasks, yet it has not been specifically optimized for the restaurant review domain. Therefore, this research focuses on investigating the impact of N-Gram feature integration and hyperparameter optimization in the fine-tuning process of the IndoBERT model to enhance the accuracy of three-class sentiment classification on restaurant review data.

The research methodology involves several main stages, beginning with the collection of customer reviews from Restoran Kampung Kecil Semarang. The data were automatically labeled using the INSET lexicon for positive, neutral, and negative sentiment categories. Two labeling approaches were applied: one with N-Gram integration and one without. The IndoBERT base-p1 model was used as the foundation, which was then fine-tuned through Grid Search hyperparameter optimization. The search space was constructed based on recommendations from Devlin et al., considering combinations of batch size, learning rate, and the number of epochs. Model performance was evaluated using accuracy and F1-Score metrics.

The results show that the use of N-Gram has a consistently positive impact on model performance, both in the pre-trained stage and after fine-tuning. The fine-tuned model with N-Gram achieved an accuracy of 0.924429 and an F1-Score of 0.924816, while the model without N-Gram achieved only 0.846834 and 0.846685, respectively. This performance improvement indicates that N-Gram is effective in capturing phrase-level context and combined sentiment polarity, which are crucial for sentiment classification tasks. Additionally, the fine-tuning process with hyperparameter optimization significantly improves the model's performance compared to the pre-trained version. This study contributes to the development of sentiment analysis systems that are more adaptive to specific domains and underscores the importance of incorporating additional linguistic features to enhance the accuracy of transformer-based models.

Keywords: *sentiment classification, IndoBERT, N-Gram, fine-tuning, hyperparameter optimization*