

ABSTRAK

Perkembangan teknologi digital telah mengubah cara manusia berkomunikasi, khususnya melalui media sosial seperti *TikTok*, *Instagram*, dan *YouTube*. *Generasi Z*, yang lahir antara tahun 1997 hingga 2012, dikenal sebagai *digital native* yang sangat aktif menggunakan media sosial sebagai sarana utama untuk berinteraksi dan mengekspresikan diri. Mereka memiliki gaya komunikasi yang khas, ditandai dengan penggunaan bahasa informal, singkatan, *emoji*, hingga ekspresi emosional yang dinamis. Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan pola bahasa *Generasi Z* dalam komentar media sosial menggunakan model *IndoBERTweet*, yaitu model *deep learning* berbasis *Natural Language Processing (NLP)* yang dioptimalkan untuk teks informal berbahasa Indonesia. Data dikumpulkan melalui proses *web scraping* dari tiga platform media sosial, kemudian dilakukan pelabelan berdasarkan kategori konten yang diasosiasikan dengan *Generasi Z*, dan divalidasi oleh pakar bahasa dan psikologi. Setelah melalui tahap *preprocessing*, data dilatih menggunakan metode *fine-tuning* terhadap model *IndoBERTweet*, dengan variasi kombinasi *hyperparameter*. Hasil pengujian menunjukkan bahwa konfigurasi terbaik diperoleh pada kombinasi *batch size* 8, *epoch* 10, dan *learning rate* $3e-5$, yang menghasilkan akurasi 89,2%, *precision* 89,2%, *recall* 89,2%, dan *F1-score* 89,2%, dengan waktu pelatihan selama 30 menit 18 detik. Temuan ini membuktikan bahwa *IndoBERTweet* efektif dalam memahami karakteristik komunikasi informal khas *Generasi Z*. Penelitian ini memberikan kontribusi nyata terhadap pengembangan teknologi *NLP* berbahasa Indonesia dan dapat diterapkan dalam pengembangan *chatbot*, sistem rekomendasi konten, analisis opini publik, serta strategi komunikasi digital yang lebih adaptif terhadap karakter *Gen Z*. Secara khusus, model ini juga dapat dimanfaatkan dalam bidang *advertising*, terutama untuk mengidentifikasi dan menyesuaikan gaya bahasa promosi yang sesuai dengan preferensi *Generasi Z*, sehingga kampanye digital yang ditujukan untuk segmen ini dapat lebih relevan, terarah, dan berdampak secara emosional.

Kata kunci: *IndoBERTweet*, *Generasi Z*, Pola Bahasa, Media Sosial.

ABSTRACT

The development of digital technology has transformed the way humans communicate, particularly through social media platforms such as TikTok, Instagram, and YouTube. Generation Z, born between 1997 and 2012, is recognized as digital natives who are highly active in using social media as their primary medium for interaction and self-expression. They demonstrate a distinctive communication style characterized by the use of informal language, abbreviations, emojis, and dynamic emotional expressions. This study aims to identify and classify the linguistic patterns of Generation Z in social media comments using IndoBERTweet, a deep learning model based on Natural Language Processing (NLP) optimized for informal Indonesian texts. Data were collected through web scraping from three social media platforms, followed by labeling based on content categories associated with Generation Z, and validated by experts in linguistics and psychology. After preprocessing, the dataset was trained through fine-tuning of the IndoBERTweet model with various hyperparameter configurations. The experimental results show that the best configuration was achieved with a batch size of 8, 10 epochs, and a learning rate of $3e-5$, yielding an accuracy of 89.2%, precision of 89.2%, recall of 89.2%, and an F1-score of 89.2%, with a training time of 30 minutes and 18 seconds. These findings demonstrate that IndoBERTweet is effective in capturing the informal communication characteristics of Generation Z. This research contributes significantly to the advancement of Indonesian NLP technology and can be applied to the development of chatbots, content recommendation systems, public opinion analysis, and digital communication strategies that are more adaptive to the characteristics of Generation Z. Specifically, this model can also be utilized in the field of advertising, particularly to identify and adjust promotional language styles that align with the preferences of Generation Z, thereby enabling digital campaigns targeted at this segment to be more relevant, focused, and emotionally impactful.

Keywords: *IndoBERTweet, Generation Z, Language Patterns, Social Media.*