

ABSTRAK

Media sosial X menjadi salah satu platform komunikasi digital yang menghasilkan teks dengan karakteristik bahasa yang beragam, termasuk penggunaan gaya komunikasi humor dan serius. Perbedaan pola bahasa tersebut menjadi tantangan dalam proses klasifikasi teks karena setiap kategori memiliki karakteristik linguistik yang berbeda. Penelitian ini bertujuan untuk mengklasifikasikan gaya komunikasi humor dan serius pada media sosial X menggunakan pendekatan *Deep Learning* serta membandingkan performa model *Convolutional Neural Network* (CNN) dan model Hibrida CNN-LSTM. Dataset yang digunakan berasal dari SmSA (*Sentiment Analysis*) pada *benchmark* IndoNLU, yang tersedia di beberapa platform publik termasuk Hugging Face, dan pada penelitian ini diakses melalui repositori resmi IndoNLU di GitHub, dengan jumlah 12.679 data teks berbahasa Indonesia. Tahapan penelitian meliputi pelabelan otomatis berbasis kata kunci, penyeimbangan kelas data, *preprocessing* teks dengan mekanisme pencegahan kebocoran label (*anti label-leakage*), tokenisasi, *padding sequence*, pembangunan model, pelatihan, dan evaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Model CNN digunakan sebagai model pembanding, sedangkan model Hibrida CNN-LSTM dikembangkan dengan menggabungkan kemampuan CNN dalam mengekstraksi fitur lokal teks dan *Bidirectional LSTM* dalam menangkap hubungan urutan kata pada data teks. Hasil penelitian menunjukkan bahwa kedua model mampu melakukan klasifikasi gaya komunikasi humor dan serius dengan baik. Model CNN memperoleh nilai *accuracy* sebesar 91,60%, *precision* sebesar 91,61%, *recall* sebesar 91,60%, dan *F1-score* sebesar 91,60%, sedangkan model Hibrida CNN-LSTM menghasilkan performa lebih tinggi dengan nilai *accuracy* sebesar 94,10%, *precision* sebesar 94,19%, *recall* sebesar 94,10%, dan *F1-score* sebesar 94,10%. Berdasarkan hasil evaluasi tersebut, model Hibrida CNN-LSTM menunjukkan kemampuan klasifikasi yang lebih optimal dibandingkan CNN tunggal, meskipun dengan konsekuensi waktu inferensi yang sedikit lebih besar (116,79 ms berbanding 84,47 ms per prediksi). Hal ini menunjukkan bahwa kombinasi CNN dan LSTM mampu menghasilkan representasi fitur teks yang lebih efektif dalam mengidentifikasi pola komunikasi humor dan serius pada media sosial X.

Kata kunci: *Natural Language Processing*, klasifikasi teks, gaya komunikasi, CNN-LSTM, media sosial X.