ABSTRAK

Saat ini, khususnya di media sosial, ujaran kebencian hampir selalu melibatkan bahasa yang menyinggung. Twitter (X) merupakan platform media sosial seperti yang lain, tetapi yang membedakannya adalah bahwa kiriman disebut tweet, dan terdapat fitur retweet untuk membagikannya kepada pengikut Anda. Meskipun populer, Twitter (X) semakin banyak digunakan untuk menyebarkan kebencian dan misinformasi karena sifatnya yang viral dan anonim. Dalam satu kalimat, ujaran kebencian dapat memiliki beberapa label yang merujuk ke beberapa topik. Studi ini mengeksplorasi efektivitas berbagai model berbasis BERT, termasuk BERT, BERT-CNN, BERT-LSTM, dan BERT-BiLSTM, untuk deteksi ujaran kebencian multi-label berdasarkan panjang teks yang berbeda. Hasil penelitian menunjukkan bahwa kinerja model bervariasi tergantung pada panjang kalimat. Pada teks panjang, model BERT-BiLSTM mencapai akurasi tertinggi sebesar 83.20% dengan skor recall dan F1 yang unggul, menunjukkan kemampuannya menangkap konteks yang kompleks dan bernuansa. BERT-CNN juga menunjukkan kinerja baik pada teks panjang dengan akurasi dan presisi yang tinggi, meskipun skor F1-nya sedikit lebih rendah dibandingkan dengan BERT-BiLSTM. Sebaliknya, model BERT dan BERT-LSTM memberikan hasil yang moderat namun kurang efektif dalam menangani konteks rinci pada teks panjang. Pada teks pendek, BERT-CNN unggul dengan akurasi tertinggi sebesar 79.80% dan skor F1 sebesar 79.10%, mengindikasikan efektivitas lapisan konvolusi dalam mengekstraksi fitur penting dari konten yang singkat. BERT-LSTM juga menunjukkan presisi dan recall yang seimbang, sementara BERT-BiLSTM memiliki recall yang kuat namun akurasinya sedikit lebih rendah, menunjukkan bahwa model ini lebih optimal untuk konteks yang lebih kaya. Hasil ini menekankan pentingnya menyesuaikan arsitektur model dengan karakteristik teks: BERT-BiLSTM optimal untuk pemahaman kontekstual yang mendalam pada teks panjang, sedangkan BERT-CNN efektif dalam mengenali fitur penting pada sampel yang lebih singkat dan ringkas.

Kata kunci: Pembelajaran Mendalam, Tugas Klasifikasi, Deteksi Ujaran Kebencian, Panjang Kalimat, Media Sosial.