

ABSTRAK

Penelitian ini bertujuan untuk mengklasifikasikan opini masyarakat terhadap fenomena pernikahan dini pada platform media sosial X menggunakan pendekatan analisis sentimen multi-kelas. Data yang digunakan sebanyak 2.008 komentar yang dikumpulkan melalui teknik *crawling* dan diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Metode klasifikasi yang digunakan adalah *Bi-LSTM* (*Bidirectional Long Short-Term Memory*), dengan representasi fitur berbasis *Word2Vec* untuk mengubah data teks menjadi vektor numerik bermakna. Sebelum proses klasifikasi, dilakukan penyeimbangan data menggunakan dua pendekatan, yaitu text augmentation dan *SMOTE*. Penelitian ini juga menguji beberapa skenario pelatihan model dengan kombinasi nilai *epoch* dan *batch size* untuk mendapatkan konfigurasi terbaik. Hasil pengujian menunjukkan bahwa model *Bi-LSTM* mampu mengklasifikasikan sentimen dengan baik, dan performa tertinggi diperoleh pada data hasil augmentasi teks dengan akurasi sebesar 90,6% menggunakan kombinasi *epoch* 25 dan *batch size* 256. Hasil ini membuktikan bahwa pendekatan *Bi-LSTM* dengan fitur *Word2Vec* efektif dalam melakukan klasifikasi sentimen multi-kelas.

Kata kunci: Analisis Sentimen, *Bi-LSTM*, *Word2Vec*, Pernikahan Dini, Media Sosial.

ABSTRACT

This study aims to classify public opinions on the phenomenon of early marriage on the social media platform X using a multi-class sentiment analysis approach. The dataset consists of 2,008 comments collected through web crawling and categorized into three sentiment classes: positive, negative, and neutral. The classification method used is Bi-LSTM (Bidirectional Long Short-Term Memory), with feature representation based on Word2Vec to convert textual data into meaningful numerical vectors. Prior to classification, data balancing was performed using two approaches: text augmentation and SMOTE. This study also evaluates several model training scenarios with various combinations of epochs and batch sizes to obtain the optimal configuration. The experimental results demonstrate that the Bi-LSTM model can effectively classify sentiment, with the best performance achieved on augmented text data, reaching an accuracy of 90.6% using 25 epochs and a batch size of 256. These results indicate that the Bi-LSTM approach with Word2Vec features is effective in performing multi-class sentiment classification.

Keywords: *Sentiment Analysis, Bi-LSTM, Word2Vec, Early Marriage, Social Media.*