ABSTRACT

IMPLEMENTATION AND EVALUATION OF RECURRENT NEURAL NETWORK (RNN) ALGORITHM ON FAKE NEWS DETECTION USING TF-IDF REPRESENTATION

By
Farrel Amroe Azhari
21/476754/PA/20600

Fake news detection has become a significant challenge in the digital information age, especially due to the complexity of massive and diverse text data. This research proposes a Natural Language Processing (NLP) based approach by comparing the performance of two *deep learning* models, namely BiLSTM (Bidirectional Long Short-Term Memory) and BiGRU (Bidirectional Gated Recurrent Unit). The Fake and Real News dataset from Gossipcop is used as the data source for experiments, with text representation performed using Term Frequency-Inverse Document Frequency (TF-IDF) and tokenization techniques to prepare the data for *deep learning*. The BiLSTM and BiGRU models were trained to classify news into two categories, namely fake news and real news. The evaluation results show that both models are able to achieve high accuracy, but have differences in terms of training time efficiency and generalization ability. The performance comparison is based on metrics such as accuracy, precision, recall, and F1-score. This research provides an in-depth insight into the advantages and disadvantages of both models in text classification tasks, as well as contributes to the development of a more effective fake news detection system.

Keywords: Fake News, Text Classification, Natural Language Processing, BiLSTM, BiGRU

INTISARI

IMPLEMENTASI DAN EVALUASI ALGORITMA RECURRENT NEURAL NETWORK (RNN) PADA DETEKSI BERITA PALSU MENGGUNAKAN REPRESENTASI TF-IDF

Oleh
Farrel Amroe Azhari
21/476754/PA/20600

Deteksi berita palsu telah menjadi tantangan signifikan dalam era informasi digital, terutama karena kompleksitas data teks yang sangat besar dan beragam. Penelitian ini mengusulkan pendekatan berbasis *Natural Language Processing* (NLP) dengan membandingkan performa dua model *deep learning*, yaitu *BiLSTM (Bidirectional Long Short-Term Memory)* dan *BiGRU (Bidirectional Gated Recurrent Unit)*. *Dataset Fake and Real News* dari Gossipcop digunakan sebagai sumber data untuk eksperimen, dengan representasi teks dilakukan menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) dan tokenisasi untuk mempersiapkan data untuk *deep learning*. Model *BiLSTM* dan *BiGRU* dilatih untuk mengklasifikasikan berita menjadi dua kategori, yaitu berita palsu dan berita asli. Hasil evaluasi menunjukkan bahwa kedua model mampu mencapai akurasi tinggi, namun memiliki perbedaan dalam hal efisiensi waktu pelatihan dan kemampuan generalisasi. Perbandingan kinerja didasarkan pada metrik seperti akurasi, presisi, *recall*, dan F1-score. Penelitian ini memberikan wawasan mendalam mengenai kelebihan dan kekurangan kedua model dalam tugas klasifikasi teks, serta kontribusi pada pengembangan sistem deteksi berita palsu yang lebih efektif.

Kata Kunci: Berita Palsu, Klasifikasi Teks, *Natural Language Processing*, *BiLSTM*, *BiGRU*