

Deteksi Berita Palsu pada Media Sosial Menggunakan Metode Convolutional Neural Network (CNN) dan Bidirectional Encoder Representations from Transformers (BERT)

Anindika Riska Intan Fauzy¹, Dr. Erwin Budi Setiawan²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹intananindika@students.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id.

Abstract

Along with the development of technology, created social media platforms facilitate human social life. Twitter is one of the most popular social media and is often used to exchange information. The news dissemination on this social media platform is real-time and complete with the news. Unfortunately, few tweets contain fake news or are often called hoaxes. The hoaxes on Twitter are very troubling to the public. Fake news or hoaxes can lead to misunderstandings in receiving information. Therefore, this study aims to develop a system that can detect hoaxes on Twitter to anticipate the spread of hoaxes that can harm related parties. The system developed uses a deep learning approach with Convolutional Neural Networks (CNN), Term Frequency-Inverse Document Frequency (TF-IDF), Bidirectional Encoder Representations from Transformers (BERT), and Global Vectors (GloVe). The results of this study show that fake news is detected by the system using the CNN method with baseline, BERT, and GloVe. The data has been adjusted to keywords proven hoaxes and spread in online media, such as Hoax or Not from Detik.com, CekFakta from Kompas.com, etc. The results show the highest accuracy of 98.57% using CNN with a split ratio of 90:10, baseline unigram-bigram, BERT, and Top10 corpus tweet+IndoNews with an increase of 4.7%.

Keywords: hoax, CNN, baseline, BERT, GloVe
