

Deteksi Hoaks Pemilu pada X Menggunakan CNN dengan Fitur Pembobotan TF-RF dan TF-IDF

Dila Adelia¹, Widi Astuti², Kemas Muslim Lhaksana³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹diladelia@students.telkomuniversity.ac.id, ²widiwdu@telkomuniversity.ac.id,

³kemasmuslim@telkomuniversity.ac.id

Abstrak

Media sosial X merupakan platform mikroblog untuk berbagi pemikiran dan tren singkat. Media sosial ini telah menjadi titik fokus untuk mengekspresikan pandangan politik. Meningkatnya keterlibatan politik di media sosial X telah memfasilitasi penyebaran ide yang cepat dan luas. Namun, media sosial X juga membawa risiko penyebaran informasi palsu dan berita bohong yang dapat memanipulasi opini publik. Mencegah berita bohong di media sosial sangat penting karena dapat memengaruhi hasil pemilu dan stabilitas sosial. Misalnya, media sosial X telah digunakan selama pemilu untuk menyebarkan berita bohong, seperti klaim palsu tentang manipulasi suara atau informasi yang menyesatkan tentang kualifikasi kandidat. Studi ini mengimplementasikan Convolutional Neural Network (CNN) karena kelebihanannya dalam mengenali pola yang kompleks dan mencapai kinerja tinggi dalam tugas-tugas seperti klasifikasi. Kumpulan data yang digunakan dalam studi ini terdiri dari 2.670 tweet. Kumpulan data dibagi menjadi tiga subset: 60% untuk pelatihan, 20% untuk pengujian, dan 20% untuk validasi. Penelitian ini juga menggunakan fitur pembobotan Term Frequency Relevance Frequency (TF-RF) dan Term Frequency Inverse Document Frequency (TF-IDF) untuk meningkatkan akurasi dalam mendeteksi berita palsu. Penelitian ini memiliki tujuan membandingkan fitur pembobotan TF-RF dan TF-IDF menggunakan metode klasifikasi CNN pada topik pemilihan umum 2024. Hasil pengujian menunjukkan bahwa TF-RF dan TF-IDF mencapai kinerja keseluruhan yang serupa, dengan TF-RF sedikit lebih unggul dalam recall dan F1-score. Pada saat yang sama, TF-IDF menunjukkan presisi yang sedikit lebih tinggi.

Kata kunci : Media Sosial X, Hoaks, Convolutional Neural Network, TF-RF, TF-IDF

Abstract

X social media is a microblogging platform for sharing brief thoughts and trends. It has become a focal point for expressing political views. The increased political engagement on X social media has facilitated the swift and extensive sharing of ideas. Still, it also brings the risk of spreading false information and hoaxes that can manipulate public opinion. Preventing fake news on social media is crucial because it can influence election outcomes and social stability. For example, X social media has been used during elections to spread hoaxes, such as false claims of vote tampering or misleading information about candidate qualifications. This study implements a Convolutional Neural Network (CNN) due to its advantages in recognizing complex patterns and achieving high performance in tasks like classification. The dataset used in this study consists of 2,670 tweets. The dataset is divided into three subsets: 60% for training, 20% for testing, and 20% for validation. It also uses Term Frequency Relevance Frequency (TF-RF) and Term Frequency Inverse Document Frequency (TF-IDF) weighting features to improve accuracy in detecting fake news. This study compares the TF-RF and TF-IDF weighting features using the CNN classification method on the topic of the 2024 election. The testing results indicate that both TF-RF and TF-IDF achieved similar overall performance, with TF-RF slightly excelling in recall and F1-score. At the same time, TF-IDF showed a marginally higher precision.

Keywords: X Social Media, Hoax, Convolutional Neural Network, TF-RF, TF-IDF
