

Deteksi Hoaks pada Media Sosial dengan *Convolutional Neural Network (CNN)* dan *Support Vector Machine (SVM)*

Manuel Benedict¹, Erwin Budi Setiawan²

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

¹manuelbenedict@student.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id

Abstrak

Berita hoaks sudah lama menjadi masalah pada masyarakat. Hal ini cukup mengkhawatirkan karena dapat berita hoaks dapat mengubah cara pandang seseorang ke hal yang tidak baik yang dampaknya merugikan banyak individu maupun kelompok masyarakat. *Machine learning* dan *deep learning* dapat diimplementasikan untuk mendeteksi berita hoaks. Contoh metode yang digunakan pada penelitian-penelitian sebelumnya adalah seperti *SVM (Support Vector Machine)* dan *CNN (Convolutional Neural Network)*. Penelitian ini mengusulkan penerapan metode *CNN* dan *SVM*. Selain itu, penelitian ini mengembangkan model *hybrid CNN-SVM*, yang menjadi keunikan penelitian ini. *Dataset* bersumber dari Twitter yang fokus pada topik Kasus Ferdy Sambo dan Tragedi Kanjuruhan yang terjadi pada tahun 2022. *Dataset* yang digunakan terbagi dua dengan rasio 90:10 (90% *data train* dan 10% *data test*). Setelah *dataset* melewati berbagai pra pemrosesan dan skenario, kedua algoritma mencapai performansi yang sangat baik. Hal ini dapat diketahui dari nilai akurasi untuk kedua metode yang berhasil ditingkatkan performansinya setelah ekstraksi dan ekspansi fitur diterapkan dengan ekstraksi fitur *TF-IDF (Term Frequency Inverse Document Frequency)*, pembobotan *unigram + bigram*, dan ekspansi fitur dengan *GloVe (Global Vector for Word Representation)*. Metode dengan ekspansi fitur *top 15* dan korpus *Tweet + Berita* mencapai performansi tertinggi untuk metode *CNN* (95,11%) dan metode dengan ekspansi fitur *top 1* dan korpus *Tweet* mencapai performansi tertinggi untuk metode *SVM* (95,95%).

Kata kunci: berita hoaks, *CNN*, *SVM*

Abstract

Hoax news has long been a problem for society. This is quite worrying because receiving hoax news can change a person's point of view to something that is not good, the impact of which is detrimental to many individuals and groups of people. Machine learning and deep learning can be implemented for detecting hoax news. Examples of methods used in previous studies are SVM (Support Vector Machine) and CNN (Convolutional Neural Network). This study proposes the application of the CNN and SVM methods. In addition, this research develops a CNN-SVM hybrid model, which is the uniqueness of this research. The dataset is sourced from Twitter which focuses on the topic of the Ferdy Sambo Case and the Kanjuruhan Tragedy that will occur in 2022. The dataset used is divided into two with a ratio of 90:10 (90% data training and 10% data testing). After the dataset has passed through various pre-processing and scenarios, both algorithms achieved very good performance. This can be seen from the accuracy values for the two methods which managed to improve their performance after feature extraction and expansion were applied with TF-IDF (Term Frequency Inverse Document Frequency) feature extraction, unigram + bigram weighting, and feature expansion, with GloVe (Global Vector for Word Representation). The method with the top 15 feature expansion and Tweet + News corpus achieves the highest performance for the CNN method (95.11%) and the method with the top 1 feature expansion and Tweet corpus achieves the highest performance for the SVM method (95.95%).

Keywords: hoax news, CNN, SVM

1. Pendahuluan

Berdasarkan Kamus Besar Bahasa Indonesia (KBBI), hoaks merupakan informasi bohong. Meskipun terdapat ciri-ciri yang mengidentifikasi berita hoaks, namun realitanya, informasi yang ada volumenya sangat besar sehingga ciri-ciri tersebut hampir tidak dapat dikenali. Oleh karena itu berita hoaks merupakan salah satu masalah yang sangat berdampak buruk bagi masyarakat. Ada yang memanfaatkan kemudahan berbagi informasi untuk mengubah informasi [1] untuk kepentingan individu maupun suatu kelompok. Hal ini berdampak sangat buruk karena berita hoaks dapat mengubah karakter dan cara berpikir orang yang mudah percaya [2]. Salah satu contohnya adalah berita hoaks mengenai vaksin Covid-19 [2] yang mengandung *chip* didalamnya. Hal ini membuat sebagian masyarakat enggan, bahkan menolak, untuk di disuntikkan vaksin Covid-19. Fenomena ini