

Identifikasi Berita Palsu Berbahasa Indonesia menggunakan Pendekatan Pola Tekstual dan Penilaian Kredibilitas Sumber Berita = Identification of Indonesian Fake News using Textual Pattern Approach and News Source Credibility Assessment

Dimas Sony Dewantara, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20523995&lokasi=lokal>

Abstrak

Berita palsu memiliki peranan yang besar terkait penyebaran informasi yang salah terhadap para pembacanya. Informasi yang salah dapat mempengaruhi persepsi atau dapat menimbulkan bias pada pemahaman hingga dapat mengubah pilihan seseorang dalam pengambilan keputusan. Kesalahan persepsi atau bias dalam pengambilan keputusan ini berpotensi menimbulkan konflik pada masyarakat. Dewasa ini, penyebaran berita palsu kian meningkat seiring berkembangnya media penyampaian informasi dalam bentuk daring atau online. Makin maraknya penyebaran berita palsu ini membuat perlunya dibangun suatu sistem yang dapat mendeteksi keberadaan berita palsu sehingga dapat meminimalisir dampak yang mungkin ditimbulkan. Permasalahan pendeteksian berita palsu ini merupakan permasalahan klasifikasi teks binomial yang mengelompokkan suatu konten berita menjadi berita asli atau berita palsu. Penelitian terkait pendeteksian berita palsu ini telah banyak dilakukan sebelumnya terutama dengan menggunakan pendekatan analisis pola tekstual pada konten berita. Namun pendekatan ini memiliki keterbatasan ketika produsen dari konten berisikan berita palsu memiliki kemampuan dalam mereplikasi pola tekstual tertentu yang menjadi ciri dari berita asli sehingga dapat mengelabui model prediksi yang digunakan. Penilaian kredibilitas sumber berita dapat dijadikan fitur tambahan selain fitur berupa pola tekstual. Sumber berita dengan kredibilitas yang baik memiliki kecenderungan dalam menghasilkan konten berita yang valid atau dapat dipercaya. Sebaliknya, sumber berita dengan kredibilitas yang buruk cenderung menghasilkan konten berita palsu atau menyesatkan. Penelitian ini menggunakan suatu sistem klasifikasi teks binomial untuk mendeteksi berita palsu melalui kombinasi metode analisis pola tekstual dengan kredibilitas sumber berita sebagai fitur tambahan. Algoritma yang digunakan penulis meliputi Multinomial Naive Bayes (MNB), Logistic Regression (LR), Support Vector Machine (SVM) dan Random Forest Decision Tree (RFDT) yang merupakan algoritma Machine Learning serta beberapa algoritma Deep Learning seperti Convolutional Neural Network (CNN), Long Short Term Memory (LSTM) serta kombinasi dari kedua algoritma tersebut. Penulis juga menggunakan metode Word Embedding Bahasa Indonesia sebagai penunjang penggunaan metode Deep Learning. Hasil dari percobaan menunjukkan klasifikasi dengan CNN pada skenario fitur pernyataan dikombinasikan dengan fitur info redaksi menghasilkan F1-Score tertinggi sebesar 0.9354 atau sekitar 93%.

.....Fake news has a big role in spreading misinformation to its readers. Incorrect information can affect perceptions or can cause bias in understanding so that it can change a person's choices in decision making. Misperceptions or biases in making these decisions have the potential to cause conflict in the community. Today, the spread of fake news is increasing along with the development of information delivery media online or online. The more widespread the spread of fake news makes it necessary to build a system that can detect fake news to minimize the impact that may be caused. The problem of detecting fake news is classifying binomial text, which groups news content into original news or fake news. Many studies related

to detecting fake news have been carried out before, especially by using a textual pattern analysis approaches on news content. However, this approaches has limitations when producers of content containing fake news have the ability to replicate certain textual patterns that characterize real news so that it can fool the predictive model used. Assessment of the credibility of news sources can be used as an additional feature in addition to features in the form of textual patterns. News sources with good credibility have a tendency to produce news content that is valid or trustworthy. On the other hand, news sources with poor credibility tend to produce false or misleading news content. This research uses a binomial text classification system to detect fake news using a combination of textual pattern analysis methods with news source credibility as an additional feature. The algorithms that will be used by the author include Multinomial Naive Bayes (MNB), Logistic Regression (LR), Support Vector Machine (SVM) and Random Forest Decision Tree (RFDT), which are Machine Learning algorithms as well as several Deep Learning algorithms such as Convolutional Neural Network (CNN) , Long Short Term Memory (LSTM) and the combination of the two algorithms. The author will also use the Indonesian Word Embedding method to support the use of the Deep Learning method. The results of the experiment show that classification with CNN in the statement feature scenario combined with the editorial info feature produces the highest F1-Score of 0.9354 or around 93%.