

Text Classification untuk Verifikasi Fakta pada Kanal Berita Bahasa Indonesia menggunakan Deep Learning dengan Ensemble Technique = Text Classification for Fact Verification for Indonesian News Channel Using Deep Learning with Ensemble Technique

Jauzak Hussaini Windiatmaja, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20513494&lokasi=lokal>

Abstrak

Sumber informasi di jejaring berita daring adalah instrumen yang memungkinkan individu membaca berita, menerbitkan berita, dan berkomunikasi. Hal ini sudah menjadi tren dalam masyarakat yang sangat mobile. Oleh karena itu, proses verifikasi fakta suatu pemberitaan menjadi sangat penting. Dengan pertimbangan tersebut, sebuah tools berbasis web service untuk verifikasi fakta menggunakan metode deep learning dengan teknik ensemble dibangun. Penggunaan teknik ensemble pada model deep learning adalah proses beberapa model pembelajaran mesin digabungkan secara strategis untuk menyelesaikan masalah menggunakan lebih dari satu model. Untuk melatih model, dibangun sebuah dataset. Dataset berisi pasangan klaim dan label. Klaim dibangun dengan data crawling di kanal berita berbahasa Indonesia. Tiga model deep learning dibangun dan dilatih menggunakan dataset yang dibuat, dengan arsitektur jaringan dan hyperparameter yang berbeda. Setelah model dilatih menggunakan dataset, ketiga model diagregasikan untuk membentuk sebuah model baru. Untuk memastikan bahwa model agregat berfungsi lebih baik daripada model tunggal, performa model deep learning ensemble dibandingkan dengan model deep learning dasar. Hasil penelitian menunjukkan bahwa model ensemble memiliki akurasi 85,18% sedangkan model tunggal memiliki akurasi 83,9%, 83,19%, dan 81,94%. Hasil ini menunjukkan bahwa model ensemble yang dibangun meningkatkan kinerja verifikasi fakta dari tiga model tunggal. Hasil penelitian juga menunjukkan bahwa metode deep learning mengungguli performa metode machine learning lain seperti naive bayes dan random forest. Untuk memvalidasi kinerja tools yang dibangun, response time dari web service diukur. Hasil pengukuran menunjukkan rata-rata response time 6.447,9 milidetik.

.....Information sources on social networks are instruments that allow individuals to read news, publish news, and communicate. This is a trend in a highly mobile society. Therefore, the process of verifying facts is very important. With these considerations, we built a web service-based tool for fact verification using deep learning methods with ensemble technique. The use of ensemble techniques in deep learning models is a process in which several machine learning models are combined to solve problems. To train the model, we created a dataset. Our dataset of Indonesian news contains pairs of claims along with labels. Claims are built by crawling data on Indonesian news channels. Three deep learning models have been built and trained using the previously created dataset with different network architectures and hyperparameters. After the model is trained, three models are aggregated to form a new model. To ensure that the aggregated model performs better than the single model, the deep learning ensemble model is compared to the single models. The results showed that the ensemble model has an accuracy of 85.18% while the single models have an accuracy of 83.9%, 83.19%, and 81.94% consecutively. These results indicate that the ensemble model built improves the fact-verification performance of the three single models. The results also show that by using the same dataset, deep learning methods outperform other machine learning methods such as naive bayes and random forest. To validate the performance of the tools we created, the response time of the web service

is measured. The measurement result shows an average response time of 6447.9 milliseconds.