

Perbandingan Machine Learning dan Deep Learning pada Klasifikasi Teks dan Analisis Sentimen terhadap Dampak Covid-19 di Indonesia pada Twitter dengan Pendekatan Multi-Label = Comparative of Machine Learning and Deep Learning on Text Classification and Sentiment Analysis on the Impact of Covid-19 in Indonesia on Twitter Using a Multi-Label Approach

Nadya Safitri, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920554612&lokasi=lokal>

Abstrak

Pemilihan metode machine learning atau deep learning menjadi suatu permasalahan dalam klasifikasi. Hal ini didapatkan dari penelitian yang menunjukkan bahwa deep learning kinerjanya lebih baik daripada machine learning, namun terdapat penelitian bahwa kedua metode tersebut kinerjanya tidak menentu tergantung dataset yang digunakan. Oleh karena itu, penelitian ini membandingkan kinerja dari machine learning dan deep learning untuk permasalahan klasifikasi teks dan analisis sentimen terhadap dampak Covid-19 di Indonesia. Hasil penelitian ini menunjukkan bahwa kinerja pada klasifikasi teks dan analisis sentimen menggunakan metode machine learning lebih baik dibandingkan dengan deep learning. Hasil penelitian mengenai klasifikasi teks menunjukkan bahwa kinerja metode machine learning yaitu Label Powerset dan Random Forest menghasilkan akurasi 77 % sedangkan kinerja metode deep learning yaitu Long Short-Term Memory (LSTM) dan Gate Recurrent Unit (GRU) menghasilkan akurasi 48%. Hasil penelitian mengenai analisis sentimen menunjukkan bahwa kinerja metode machine learning yaitu Label Powerset dan Random Forest menghasilkan akurasi 63 % sedangkan kinerja metode deep learning yaitu Long Short-Term Memory (LSTM) dan Gate Recurrent Unit (GRU) menghasilkan akurasi 55% dan 54%. Keseimbangan jumlah label pada semua label mempengaruhi hasil dari klasifikasi. Oleh karena itu, disarankan untuk menggunakan metode untuk menyeimbangkan jumlah label yang digunakan untuk klasifikasi.

.....The choice of machine learning or deep learning methods becomes a problem in classification. This is obtained from research which shows that deep learning performs better than machine learning, but there is research that the two methods perform erratically depending on the dataset used. Therefore, this study compares the performance of machine learning and deep learning for text classification problems and sentiment analysis on the impact of Covid-19 in Indonesia. The results of this study indicate that the performance of text classification and sentiment analysis using machine learning methods is better than deep learning. The results of research on text classification show that the performance of machine learning methods, namely Label Power and Random Forest, produces an accuracy of 77%, while the performance of deep learning methods, namely Long Short-Term Memory (LSTM) and Gate Recurrent Unit (GRU), produces an accuracy of 48%. The results of the research on sentiment analysis show that the performance of machine learning methods, namely Label Power and Random Forest, produces an accuracy of 63%, while the performance of deep learning methods, namely Long Short-Term Memory (LSTM) and Gate Recurrent Unit (GRU), produces 55% and 54% accuracy. The balance of the number of labels on all labels affects the results of the classification. Therefore, it is advisable to use a method to balance the number of labels used for classification.