

Deteksi Malicious Account pada Akun Twitter Indonesia Berbasis Tweet = Malicious Accounts Detection on Indonesian Twitter Accounts using Tweets

Latifah Al Haura, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920530066&lokasi=lokal>

Abstrak

Penipuan dan bahkan pencurian informasi saat ini kerap terjadi di media sosial melalui unggahan pengguna yang tidak bertanggung jawab berupa status, tweet, ataupun pesan Spam yang berisi tautan-tautan yang berbahaya. Hal ini tidak terlepas dari keberadaan akun-akun jahat yang sudah sangat meresahkan dan mengganggu keamanan dan kenyamanan pengguna media sosial. Oleh karena itu, penelitian ini bertujuan untuk menggunakan fitur dari tweet (teks) dalam mendeteksi Malicious Account (akun jahat) di Twitter pengguna Indonesia. Terdapat dua metode ekstraksi fitur teks yang digunakan dan dibandingkan dalam penelitian ini yaitu Word2Vec dan FastText. Selain itu, penelitian ini juga membahas perbandingan antara metode Machine Learning dan Deep Learning dalam mengklasifikasi pengguna atau akun berdasarkan fitur dari tweet tersebut. Algoritma Machine Learning yang digunakan di antaranya adalah Logistic Regression, Decision Tree, dan Random Forest sedangkan algoritma Deep Learning yang digunakan yaitu Long Short-Term Memory (LSTM). Hasil dari keseluruhan skenario pengujian menunjukkan bahwa performa rata-rata yang dihasilkan metode ekstraksi fitur Word2Vec lebih unggul dibandingkan dengan FastText yang memiliki nilai F1-Score sebesar 74% dan metode klasifikasi Random Forest lebih unggul dibandingkan dengan tiga metode lainnya yang mana memiliki nilai F1-Score sebesar 82%. Sedangkan performa terbaik untuk kombinasi antara metode ekstraksi fitur dan metode klasifikasi terbaik yaitu gabungan antara Pre-trained Word2Vec dan LSTM dengan nilai F1-Score sebesar 84%.

.....Fraud and even theft of information nowadays often occur on social media through irresponsible user uploads in the form of statuses, tweets, or spam messages containing dangerous links. This is inseparable from the existence of Malicious Accounts that have been very disturbing and disturbing the comfort of users and the comfort of social media users. Therefore, this study aims to use the feature of tweets (text) in detecting Malicious Accounts on Indonesian Twitter users. There are two text feature extraction methods used and compared in this study, namely Word2Vec and FastText. In addition, this study also discusses the comparison between Machine Learning and Deep Learning methods in classifying users or accounts based on the features of the tweet. The Machine Learning algorithm used is Logistic Regression, Decision Tree, and Random Forest, while the Deep Learning algorithm used is Long Short-Term Memory (LSTM). The results of all test scenarios show that the average performance of the Word2Vec feature extraction method is higher than FastText with an F1-Score value of 74% and the Random Forest classification method is higher than the other three methods which have an F1-Score value of 82%. While the best performance for the combination of feature extraction method and the best classification method is the combination of Pre-trained Word2Vec and LSTM with an F1-Score value of 84%.