

ABSTRAK

Deteksi Hoaks Pada Twitter

Titi Widaretna

Pembimbing 1 : Jimmy Tirtawangsa, Ph.D

Pembimbing 2 : Dr. Ade Romadhony

Hoax adalah informasi menyesatkan yang bisa sangat berbahaya. Penelitian ini bertujuan untuk mendeteksi tweet sebagai hoax atau non-hoax berdasarkan konteks tweet dengan menggunakan pendekatan semantik dan teknik klasifikasi. Untuk menemukan konteks pada tweet, Doc2Vec digunakan untuk menghasilkan representasi vektor semantik yang dapat menangkap konteks tersebut. Vektor semantik yang dihasilkan oleh Doc2Vec akan digunakan sebagai input dalam proses klasifikasi. Beberapa algoritma pengklasifikasi digunakan untuk menemukan algoritma terbaik agar dapat mendeteksi hoax secara akurat.

Pendeteksian hoax di Twitter fokus pada pengukuran nilai akurasi dan presisi, dimana tingkat kebenaran dalam mendeteksi hoax dan non-hoax harus tinggi (True Positive dan True Negative), sedangkan tingkat kesalahan untuk mendeteksi hoax sebagai non-hoax haruslah rendah (False Positive). Hasil eksperimen dari metode yang diusulkan, hasil terbaik diperoleh dengan menggunakan kombinasi Doc2Vec untuk model representasi tweet dan SVM sebagai pengklasifikasi dengan akurasi 93,02% dan presisi 93,02%.

Kata Kunci: Deteksi Hoax, Doc2Vec, Natural Language Preprocessing