

Deteksi Konten *Hoax* Berbahasa Indonesia di Twitter Menggunakan Fitur Ekspansi dengan *Word2Vec*

Friskadini Ismayanti¹, Erwin Budi Setiawan²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹friskadiniism@students.telkomuniversity.ac.id, ²erwinbudisetiawan@telkomuniversity.ac.id,

Abstrak

Penggunaan Internet di Indonesia semakin terus meningkat, dikarenakan semakin banyak orang yang dapat mengakses Internet maka semakin banyak juga penyalahgunaan dalam Internet itu sendiri khususnya media sosial. Perlu untuk dapat mendeteksi adanya *hoax* di media sosial seperti *Twitter*. *Twitter* sebagai salah satu media sosial memiliki peran yang signifikan dalam penyebaran informasi, semua orang dapat memberikan respons terhadap informasi yang didapat. Dalam isi sebuah *tweet* memungkinkan terjadinya ketidakcocokan kosakata. Oleh karena itu pada penelitian ini dilakukan penerapan metode *Feature Expansion Word2Vec* untuk mengatasi terjadinya ketidakcocokan kosakata. Penelitian ini melakukan pengembangan dan perbandingan sistem klasifikasi *hoax* *Twitter* menggunakan metode *Feature Expansion Word2Vec* dengan algoritma klasifikasi *Logistic Regression*, *Support Vector Machine* (SVM), *Random Forest* dan sistem tanpa metode *Feature Expansion Word2Vec*. Hasil dari penelitian ini, metode *Feature Expansion Word2Vec* pada algoritma klasifikasi *Random Forest* berhasil meningkatkan akurasi sistem sebesar 1,46% dengan nilai akurasi sebesar 89,53%.

Kata kunci : *Word2Vec*, *Random Forest*, *Feature Expansion*, *Twitter*, *klasifikasi hoax*
