

## Abstrak

Twitter sebagai salah satu media sosial populer untuk menyalurkan opini berupa kritik dan saran. Kritik dapat dikategorikan ujaran kebencian jika kritik tersebut menyerang sesuatu yang ditargetkan. Dengan adanya batasan 280 karakter dalam suatu tweet, seringkali terjadi ketidakcocokan kosakata dalam tweet karena penyingkatan tulisan. Hal ini dapat diatasi dengan *word embedding* (penyematan kata). Penelitian ini memanfaatkan ekspansi fitur untuk mereduksi ketidakcocokan kosakata dalam ujaran kebencian di Twitter yang mengandung Bahasa Indonesia dengan menggunakan *Global Vectors* (GloVe). Selanjutnya dilakukan seleksi fitur terkait dengan model terbaik menggunakan algoritma *Logistic Regression* (LR), *Random Forest* (RF), dan *Artificial Neural Network* (ANN). Hasil penelitian menunjukkan ekspansi fitur cukup efektif untuk meningkatkan performa dengan tingkat akurasi mencapai 88,59% pada model RF sebanyak 5.000 fitur dengan kombinasi TF-IDF korpus Tweet. Kenaikan akurasi mencapai 2,37% oleh classifier ANN dengan Top 10 dari korpus Tweet, IndoNews dibandingkan dengan *baseline* yang telah ditetapkan.

**Kata kunci:** ujaran kebencian, GloVe, ekspansi fitur