

Abstrak

Media sosial berkembang secara eksponensial sejak tahun 2004 sebagai sebuah wadah untuk berdiskusi dan bertukar pendapat. Perkembangan ini mendapatkan antusiasme yang baik oleh warga negara Indonesia. Di mana dari 170 juta jiwa penduduk negara Indonesia menjadi pengguna aktif sosial media. Namun tujuan dari pembuatan media sosial sering sekali disalahgunakan untuk menyebarkan komentar beracun, seperti menyebarkan kebencian, pornografi, radikalisme, SARA, dan masih banyak lagi. Sehingga, akhir- akhir ini analisis sentimen terhadap komentar beracun pada media sosial sedang marak dilakukan. Maka dari itu pada penelitian ini akan dilakukan analisis sentimen terhadap komentar beracun pada media sosial untuk memilah komentar yang beracun atau tidak beracun berdasarkan label atau *multilabel*. Pada penelitian ini, digunakan metode Word2Vec sebagai ekstraksi fitur. Word2Vec sebagai ekstraksi fitur telah banyak diterapkan pada penelitian *Natural Language Processing* (NLP) dan menunjukkan dampak potensial yang tinggi pada kinerja analisis sentimen. Kemudian dilakukan klasifikasi menggunakan metode *Support Vector Machine* (SVM) untuk menemukan hasil terbaik. Hasil optimum yang dihasilkan dari beberapa pengujian yang dilakukan menunjukkan nilai F1-Score tertinggi sebesar 73,69% data terklasifikasi benar menggunakan Word2Vec sebagai ekstraksi fitur dan tanpa menggunakan *stemming* pada tahap *preprocessing*.

Kata kunci : Komentar beracun, Analisis sentimen, *Multilabel*, Word2Vec, *Support Vector Machine* (SVM).