

Abstrak

Sosial media adalah media online yang penggunaannya dapat berpartisipasi, berbagi, dan membuat isi seperti forum, blog, jejaring sosial, wiki dan dunia virtual. Salah satu media sosial dengan platform komunikasi yang banyak digunakan adalah Twitter, dengan berbagai “tweet” dengan isi yang berbeda-beda dilontarkan setiap harinya, seperti topik terkait politik, sosial, pemerintah, pelayanan, dll, dengan opini dari segi positif, netral, dan negatif. Kalimat “tweet” dalam Twitter mempunyai batasan yaitu sebanyak 280 karakter, oleh karena itu menyebabkan permasalahan seperti ketidakcocokan kosakata. Oleh karena itu, penelitian ini menerapkan Word2Vec untuk mengatasi permasalahan tersebut dan metode klasifikasi *Support Vector Machine* (SVM). Penelitian ini juga menggunakan ekstraksi fitur TF-IDF digunakan sebagai pembobotan kosakata. Teknik *Synthetic Minority Oversampling Technique* atau SMOTE digunakan untuk mengatasi *data imbalanced* atau ketidakseimbangan data, dan dengan perbandingan ekstensi SMOTE Tomek Links, dan SMOTE ENN. Hasil dari penelitian ini didapatkan hasil performansi *F1-score* tertinggi dengan menggunakan klasifikasi Support Vector Machine dan teknik *imbalance data* SMOTE + ENN dengan ekstraksi fitur TF-IDF, seleksi fitur, dan ekspansi fitur Word2Vec pada korpus twitter + berita pada aspek sinyal dengan peningkatan 28.69% dengan *F1-score* 98.58%, dan pada aspek pelayanan dengan peningkatan 27.79% dengan *F1-score* 95.44%.

Kata kunci : Analisis Sentimen berdasarkan Aspek, Support Vector Machine, Word2vec, Twitter, Telkomsel