

ABSTRAK

Investasi merupakan pengumpulan aset untuk memperoleh keuntungan di masa depan. Dalam beberapa tahun terakhir, Indonesia mengalami peningkatan signifikan dalam bidang investasi digital, sehingga banyak bermunculan platform investasi digital. Namun, banyak masyarakat yang kebingungan dalam memilih platform investasi digital yang tepat. Oleh karena itu, penelitian ini mengusulkan pendekatan untuk menganalisis aplikasi investasi digital, salah satunya Bareksa, dengan menggunakan pendekatan lexicon-based untuk pelabelan, Support Vector Machine (SVM) dengan kernel linear, seleksi fitur Chi-Square, dan metode Synthetic Minority Over-sampling Technique (SMOTE). Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Bareksa. Salah satu tantangan utama dalam analisis sentimen adalah banyaknya fitur tidak relevan (noisy features) dalam data teks yang dapat menurunkan kinerja model. Untuk mengatasi masalah ini, digunakan seleksi fitur Chi-Square untuk menyaring fitur-fitur yang tidak signifikan. Selain itu, penelitian ini juga menghadapi masalah ketidakseimbangan data (imbalance data), di mana ulasan positif lebih banyak daripada ulasan negatif. Untuk mengatasi ketidakseimbangan ini, diterapkan metode SMOTE. Hasil penelitian menunjukkan bahwa penggunaan seleksi fitur Chi-Square, SMOTE, dan SVM memberikan hasil yang optimal dengan akurasi 93,46%, presisi 89,42%, dan recall 89,30% menggunakan k-fold 7 dan seleksi 90% fitur. Sebagai perbandingan, algoritma SVM tanpa penerapan seleksi fitur dan SMOTE hanya memberikan akurasi tertinggi sebesar 90,15%, presisi 90,13%, dan recall 90,15%. Temuan ini menunjukkan bahwa kombinasi pendekatan ini efektif dalam meningkatkan kinerja analisis sentimen ulasan aplikasi Bareksa.

Kata kunci : Analisis Sentimen, TF-IDF, *Lexicon Based*, *Chi Square*, *Support Vector Machine*, *smote*