

ABSTRAK

Pulau Madura memiliki berbagai destinasi wisata yang menawarkan keindahan alam memukau, namun mengalami penurunan jumlah wisatawan dan ulasan negatif terkait fasilitas dan pelayanan. Penelitian ini bertujuan menganalisis persepsi wisatawan terhadap destinasi wisata di Madura melalui analisis sentimen. Metode yang digunakan dalam penelitian ini adalah *Random Forest* yang unggul dalam menangani data besar dan mengatasi data hilang. Namun, kinerjanya dapat menurun apabila jumlah fitur yang sangat besar sehingga dapat menyebabkan penurunan performa pada model. Oleh karena itu, digunakan seleksi fitur dengan metode *Information Gain* dan *Chi-Square* melalui pendekatan *ensemble* untuk mereduksi dimensi fitur. Hasil pengujian dari metode *Random Forest* tanpa seleksi fitur, didapatkan akurasi terbaik, yakni sebesar sebesar 77,26% dengan kinerja waktu proses 3,38 detik pada pembagian data 80:20. Kemudian, pengujian dari masing-masing seleksi fitur diperoleh akurasi tertinggi pada seleksi fitur *Information Gain* dengan mencapai 79,90% dengan kinerja waktu proses 4,38 detik (data 90:10, *threshold* 0,001). Pendekatan *ensemble* seleksi fitur dengan *union* menghasilkan akurasi 78,35% dengan kinerja waktu proses 4,70 detik (data 90:10). Selanjutnya, penggunaan *information gain* dengan SMOTE 2 kelas menghasilkan akurasi tertinggi, yaitu 81,47% dengan kinerja waktu proses 1,97 detik (data 80:20, *threshold* 0,005).

Kata Kunci: Analisis Sentimen, *Ensemble*, *Information Gain*, *Random Forest*, Seleksi Fitur.