

ANALISIS SENTIMEN PADA ULASAN WISATA PANTAI SEMBILAN SUMENEP MENGGUNAKAN *K-NEAREST NEIGHBOR* DAN SELEKSI FITUR *INFORMATION GAIN*

ABSTRAK

Kota Sumenep merupakan salah satu daerah yang memiliki berbagai macam obyek wisata, salah satu obyek wisata yang sudah menjadi ikon wisata Sumenep adalah Pantai Sembilan. Banyaknya pengunjung yang datang, tidak sedikit yang memberikan ulasan di *google maps*. Ulasan pengunjung terhadap obyek wisata ini dapat dijadikan sebagai masukan bagi peningkatan layanan dan pengembangan obyek wisata dengan melakukan suatu analisis sentimen. Metode yang digunakan untuk klasifikasi analisis sentimen pada penelitian ini yaitu *K-Nearest Neighbor* dengan seleksi fitur *Information gain*. Metode *K-Nearest Neighbor* cukup baik dalam mengklasifikasikan data teks, akan tetapi mempunyai kelemahan ketika bekerja dengan jumlah fitur yang banyak. Adapun tujuan dari penelitian ini adalah menerapkan seleksi fitur *information gain* untuk mengurangi jumlah fitur sehingga dapat meningkatkan performa *K-Nearest Neighbor*. Data yang digunakan pada penelitian ini berjumlah 1318 data ulasan. Kemudian dilakukan *text preprocessing* yang meliputi *casefolding*, *cleansing*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming* yang selanjutnya dilakukan pembobotan kata menggunakan TF-IDF. Hasil yang diperoleh dari pengujian model klafikasi *K-Nearest Neighbor* didapatkan akurasi terbaik sebesar 98,36% pada nilai k yaitu 1 tanpa seleksi fitur dan 97,54% pada nilai k yaitu 3 menggunakan seleksi fitur *information gain* dengan batas *threshold* 0,0008

Kata Kunci : Analisis sentimen, wisata, Pantai Sembilan, *KNN*, *information gain*, *google maps*.