

Implementasi Algoritma *Naïve Bayes* Terhadap Analisis Sentimen Kepuasan Pengunjung Wisata Di Kabupaten Bangkalan

Ayussy Rahma Aulia

(190441100053)

Pembimbing 1 : Dr. Wahyudi Setiawan, S.Kom.,M.Kom.

Pembimbing 2 : Doni Abdul Fatah, S.Kom.,M.Kom.

ABSTRAK

Analisis sentimen kepuasan pengunjung wisata merupakan salah satu upaya yang dilakukan untuk mempermudah pengambilan keputusan bagi pengelola tempat wisata di Kabupaten Bangkalan dalam mengembangkan tempat wisata. Pada penelitian ini dilakukan analisis sentimen kepuasan pengunjung pada ulasan wisata di Kabupaten Bangkalan menggunakan *Naive bayes* dengan pembobotan kata TF-IDF disertai seleksi fitur *Information gain*. Data yang digunakan merupakan ulasan dari *Google Maps* yang diambil melalui *web scraping* menggunakan tools *WebHarvy*. Hasil pelabelan kelas sentimen pada data ulasan pengunjung wisata di Kabupaten Bangkalan diperoleh 3649 ulasan dengan jumlah ulasan positif sebanyak 2583, negatif sebanyak 275 ulasan dan netral sebanyak 457 ulasan. *Imbalance class* pada dataset ditangani dengan menggunakan teknik SMOTE. Hasil analisa menunjukan penggunaan seleksi fitur *information gain* dan teknik SMOTE pada model *Naïve bayes* dengan *threshold* 0.0001 dinilai lebih efektif karena menghasilkan nilai akurasi sebesar 83.00%, presisi 83.44%, recall 82,59% dan f1-score 82.53%. Hal ini menunjukan bahwa penggunaan seleksi fitur *information gain* dan teknik SMOTE memiliki performa yang cukup baik dalam melakukan klasifikasi pada data ulasan pengunjung wisata di Kabupaten Bangkalan

Kata kunci: Analisis Sentimen, *Naïve Bayes*, *TF-IDF*, *Information Gain*, SMOTE, Wisata, Kabupaten Bangkalan.

***Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Tourist Satisfaction in
Bangkalan Regency***

Ayussy Rahma Aulia

(190441100053)

Pembimbing 1 : Dr. Wahyudi Setiawan, S.Kom.,M.Kom.

Pembimbing 2 : Doni Abdul Fatah, S.Kom.,M.Kom.

ABSTRACT

Sentiment analysis of tourist satisfaction is one of the efforts carried out to facilitate decision-making for tourism site managers in Bangkalan Regency in developing tourist destinations. In this research, sentiment analysis of visitor satisfaction in tourist reviews in Bangkalan Regency was conducted using Naive Bayes with TF-IDF word weighting and accompanied by Information Gain feature selection. The data used consisted of reviews from Google Maps, obtained through web scraping using WebHarvy tools. The labeled results of sentiment classes in the tourist reviews data of Bangkalan Regency obtained 3649 reviews, with 2583 positive reviews, 275 negative reviews, and 457 neutral reviews. The dataset's class imbalance was addressed using the SMOTE technique. The analysis results show that the use of Information Gain feature selection and SMOTE technique in the Naive Bayes model with a threshold of 0.0001 is considered more effective as it resulted in an accuracy of 83.00%, precision of 83.44%, recall of 82.59%, and f1-score of 82.53%. This indicates that the use of Information Gain feature selection and SMOTE technique has performed well in classifying tourist reviews data in Bangkalan Regency.

Keywords: *Sentiment Analysis, Naive Bayes, TF-IDF, Information Gain, SMOTE, Tourism, Bangkalan Regency.*