

ABSTRAK

Pelayanan administrasi kependudukan secara *online* menjadi prioritas pemerintah daerah untuk meningkatkan efisiensi dan kepuasan masyarakat. Kabupaten Cilacap meluncurkan aplikasi "DOLAN TELUK PENYU" sebagai sarana layanan publik digital, namun keberhasilannya bergantung pada penerimaan dan opini pengguna sehingga diperlukan analisis mendalam terhadap ulasan pengguna secara objektif. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap aplikasi "DOLAN TELUK PENYU" serta membandingkan kinerja algoritma *Naive Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest* sebelum dan sesudah penerapan *SMOTE (Synthetic Minority Over-sampling Technique)* untuk mengatasi ketidakseimbangan kelas data. Data penelitian berupa 1.198 komentar pengguna aplikasi pada *platform YouTube* yang setelah proses pembersihan dan penghapusan data duplikat menghasilkan 1.163 data siap olah. Tahapan penelitian meliputi *crawling* dan *scraping data*, *preprocessing* teks (*cleaning*, *case folding*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming*), pelabelan sentimen menggunakan pendekatan *lexicon based*, ekstraksi fitur dengan *TF-IDF*, klasifikasi, hingga evaluasi model menggunakan *confusion matrix*. Hasil pelabelan menunjukkan distribusi sentimen Netral 440 data, Positif 395 data, dan Negatif 328 data. Sebelum *SMOTE*, *Random Forest* memperoleh akurasi tertinggi sebesar 78,97%, diikuti *SVM* 72,53% dan *Naive Bayes* 70,39%. Setelah *SMOTE* diterapkan, akurasi *Random Forest* meningkat menjadi 79,17% dan *SVM* meningkat menjadi 77,27%, sedangkan *Naive Bayes* justru menurun menjadi 63,64%. *Random Forest* terbukti sebagai model yang paling stabil dan akurat, sehingga hasil penelitian ini dapat menjadi dasar evaluasi bagi Disdukcapil Kabupaten Cilacap dalam meningkatkan kualitas aplikasi dan pelayanan publik berbasis *e-government*.

Kata kunci: Dolan Teluk Peny, Analisis Sentimen, *SMOTE*, *Naive Bayes*, *Random Forest*.

ABSTRACT

Online population administration services have become a priority for local governments to improve efficiency and public satisfaction. Cilacap Regency launched the "DOLAN TELUK PENYU" application as a digital public service platform, yet its success depends on user acceptance and opinion, requiring an in-depth and objective analysis of user reviews. This research aims to analyze public sentiment toward the "DOLAN TELUK PENYU" application and to compare the performance of the Naive Bayes, Support Vector Machine (SVM), and Random Forest algorithms before and after applying SMOTE (Synthetic Minority Over-sampling Technique) to address class imbalance. The research data consisted of 1,198 user comments collected from YouTube, which, after data cleaning and duplicate removal, resulted in 1,163 ready-to-process comments. The research stages included data crawling and scraping, text preprocessing (cleaning, case folding, normalization, tokenizing, stopword removal, and stemming), sentiment labeling using a lexicon-based approach, feature extraction with TF-IDF, classification, and model evaluation using a confusion matrix. The labeling results showed a sentiment distribution of 440 Neutral, 395 Positive, and 328 Negative comments. Before SMOTE, Random Forest achieved the highest accuracy of 78.97%, followed by SVM at 72.53% and Naive Bayes at 70.39%. After SMOTE was applied, the accuracy of Random Forest increased to 79.17% and SVM increased to 77.27%, while Naive Bayes decreased to 63.64%. Random Forest proved to be the most stable and accurate model, so the results of this research can serve as a basis for evaluation by the Disdukcapil of Cilacap Regency in improving the quality of the application and e-government-based public services.

Keywords: Dolan Teluk Penyru, Sentiment Analysis, SMOTE, Naive Bayes, Random Forest.