

PERBANDINGAN ANALISIS SENTIMEN JUDUL BERITA PARIWISATA BALI MENGGUNAKAN SVM (*SUPPORT VECTOR MACHINE*)

Oleh

I Putu Soma Darmayasa, NIM 1915101017

Jurusan Teknik Informatika

ABSTRAK

Penelitian ini memberikan kontribusi dalam pemahaman lebih baik tren-tren dan isu-isu yang sedang terjadi dalam industri pariwisata Bali melalui analisis sentimen judul berita. Judul berita tentang pariwisata Bali yang dipublikasikan oleh portal media online memiliki makna atau kesan positif, netral. Tujuan dari penelitian ini adalah mengembangkan sebuah model *machine learning* dengan membandingkan seleksi fitur yang dapat mengklasifikasi judul berita pariwisata Bali serta menganalisis berdasarkan sentimen yang terkandung didalamnya. Metode seleksi fitur yang digunakan adalah *Chi-Square* dan *Information Gain* serta model *machine learning* yang digunakan adalah SVM. Data judul berita yang digunakan diambil dari portal media online seperti Detik.com, Kompas.com, dan Balipost.com yang dilabeli secara manual sesuai dengan sentiment yang terdapat pada judul berita. Pembagian dataset dalam penelitian ini adalah 90:10. 90% untuk data training (*training proses*) dengan menggunakan *K-Fold Cross Validation* dan 10% untuk data testing atau pengujian. Hasil dari penelitian menunjukkan bahwa model SVM dengan seleksi fitur *Chi-Square* menghasilkan akurasi sebesar 88,88% model dan SVM dengan seleksi *Information Gain* menghasilkan akurasi sebesar 86,65% untuk testing data atau pengujian.

Kata Kunci: analisis sentimen, *chi-square*, judul berita, klasifikasi, *latent dirchlect allocation*, *information gain*, *support-vector machine*

**SENTIMENT ANALYSIS COMPARISON OF BALI TOURISM NEWS
HEADLINES USING SVM (SUPPORT VECTOR MACHINE)**

By

I Putu Soma Darmayasa, NIM 1915101017

Computer Science Major

ABSTRACT

This research contributes to a better understanding of trends and issues occurring in the Bali tourism industry through sentiment analysis of news headline titles. The title of news articles on tourism in Bali published by online media portals carries positive or neutral connotations. The aim of this research is to develop a machine learning model by comparing feature selection methods capable of classifying Bali tourism news headlines and analyzing the sentiment contained therein. The feature selection methods employed are Chi-Square and Information Gain, and the machine learning model used is SVM. The news headline data used are sourced from online media portals such as Detik.com, Kompas.com, and Balipost.com, manually labeled according to the sentiment present in the headlines. The dataset distribution proportion in this study is 90% for training data (training process) using K-Fold Cross Validation and 10% for testing data. The results of the research indicate that the SVM model with Chi-Square feature selection achieves an accuracy of 88.88%, and the SVM model with Information Gain feature selection achieves an accuracy of 86.65% for testing data.

Keywords: *classifications, chi-square, headline news, information gain, latent dirchlect allocation sentiment analysis, support-vector machine*