

ABSTRAK

Hendrayana, I Gede (2023), Komparasi Metode Support Vector Machine, K-Nearest Neighbor Dan Naïve Bayes Classifier Pada Analisis Sentimen Wisata Alam Jatiluwih. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I : Prof. Dr. Dewa Gede Hendra Divayana, S.Kom., M.Kom. dan Pembimbing II : Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.

Kata-kata kunci: Kata Kunci: Support Vector Machine, K-Nearest Neighbour, Naïve Bayes Classifier, Analisis Sentimen, Wisata Alam

Keindahan Bali menciptakan munculnya banyak komentar bahwa tidak lengkap berwisata ke Indonesia tanpa ke Bali sehingga banyak orang yang menyebut Bali sebagai surga wisata. Pada saat ini, wisata alam akan menjadi pilihan utama masyarakat untuk pergi liburan karena ingin menikmati udara yang segar dan keindahan alam. Sejalan dengan itu, pengembangan kawasan wisata alam yang berada di daerah diperlukan untuk memberikan sarana rekreasi ke masyarakat sehingga terjadi pemerataan pembangunan, peningkatan ekonomi masyarakat daerah dan tidak terjadi penumpukan wisata pada daerah tertentu. Untuk mendukung hal tersebut, kepuasan dan perspektif wisatawan sangat penting penerapannya namun sebagian besar masih menerapkan survei untuk mendapatkan masukan dan perspektif wisatawan. Pendekatan berbasis survei memiliki kekurangan seperti biaya operasional, potensi data ganda, hanya mencakup aspek tujuan yang telah ditentukan sebelumnya dan kurang komprehensif. Salah satu solusi yang diterapkan adalah melakukan analisis sentimen pada objek wisata alam dengan mengklasifikasikan komentar positif dan negatif menggunakan dataset komentar wisatawan. Pada penelitian ini menggunakan objek wisata alam Jatiluwih sebagai dataset komentar wisatawan dan membandingkan performa metode Support Vector Machine (SVM), K-Nearest Neighbor (K-NN) Naïve Bayes Classifier (NBC). Fokus analisis sentimen panelitian ini pada komentar yang berhubungan dengan kriteria Objek Daya Tarik Wisata Alam menurut Direktorat Jenderal Perlindungan Hutan dan Konservasi Alam tahun 2003 adalah daya tarik wisata, potensi pasar, aksesibilitas, kondisi lingkungan sosial ekonomi, pelayanan masyarakat, kondisi iklim, akomodasi, prasarana dan sarana penunjang, tersedianya air bersih dan keamanan. Penelitian ini bertujuan untuk memberikan analisis yang komprehensif perbandingan mengenai performa metode SVM, K-NN dan NBC menggunakan *confusion matrix* yang meliputi *accuracy*, *precision*, *recall* dan *f-measure*. Hasil penelitian ini menunjukkan metode K-NN dengan nilai K 11 menghasilkan *accuracy* sebesar 93.24%, *precision* sebesar 93.23%, *recall* sebesar 100%, *f-measure* sebesar 96.24% yang lebih baik daripada metode SVM dan KNN. Diharapkan untuk penelitian selanjutnya dapat melibatkan para ahli bahasa, metode pra pengolahan data yang berbeda dan optimasi metode klasifikasi sehingga dapat meningkatkan dan memberikan perspektif yang berbeda hasil klasifikasi tersebut.

ABSTRACT

Hendrayana, I Gede (2023), Comparison of Support Vector Machine, K-Nearest Neighbor, and Naïve Bayes Classifier Methods in the Sentiment Analysis of Jatiluwih Nature Tourism. Thesis, Computer Science, Postgraduate Program, Universitas Pendidikan Ganesha.

This thesis has been approved and examined by Supervisor I : Prof. Dr. Dewa Gede Hendra Divayana, S.Kom., M.Kom. and Supervisor II : Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.

Keywords: Support Vector Machine, K-Nearest Neighbour, Naïve Bayes Classifier, Sentiment Analysis, Natural Tourism

The beauty of Bali has generated many comments that traveling to Indonesia would be incomplete without visiting Bali, thus many people refer to Bali as a tourist paradise. Currently, natural tourism has become the primary choice for people to go on vacation because they want to enjoy fresh air and the beauty of nature. In line with this, the development of natural tourism areas in various regions is necessary to provide recreational facilities to the public, leading to more balanced development, improved local economies, and avoiding an overwhelming concentration of tourism in specific areas. To support this, the satisfaction and perspectives of tourists are crucial in its implementation, but mostly still obtained through surveys to gather their feedback and viewpoints. Survey-based approaches have drawbacks, such as operational costs, potential duplicate data, limited coverage of predefined aspects, and lack of comprehensiveness. One solution applied is to perform sentiment analysis on comments related to natural tourism objects, classifying them as positive or negative, using a dataset of tourist comments. This research uses Jatiluwih, a natural tourism object, as the dataset for tourist comments and compares the performance of the Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), and Naïve Bayes Classifier (NBC) methods. The focus of this sentiment analysis is on comments related to the criteria of Natural Tourist Attractions according to the Directorate General of Forest Protection and Nature Conservation in 2003. These criteria include tourism appeal, market potential, accessibility, socio-economic environmental conditions, community services, climatic conditions, accommodation, infrastructure and support facilities, availability of clean water, and security. The aim of this study is to provide a comprehensive analysis comparing the performance of SVM, K-NN, and NBC methods using a confusion matrix that includes accuracy, precision, recall, and f-measure. The results of this research show that the K-NN method with a value of K equal to 11 achieved an accuracy of 93.24%, precision of 93.23%, recall of 100%, and an f-measure of 96.24%, which is better than the SVM and KNN methods. For further research, it is hoped that language experts can be involved, different data preprocessing methods can be employed, and classification methods can be optimized to enhance and provide different perspectives on the classification results.