

## ABSTRAK

Daerah Istimewa Yogyakarta sebagai salah satu provinsi besar di Indonesia memiliki banyak *coffee shop* yang menawarkan beragam suasana, jenis kopi, dan layanan unik. Hal ini dapat dilihat dari banyaknya pilihan *coffee shop* di Daerah Istimewa Yogyakarta yang menyajikan kopi dengan cita rasa khas. Ulasan pelanggan memiliki peran penting dalam membentuk citra dan reputasi sebuah *coffee shop*, dan banyak orang mengandalkan *review* dari pengguna *Google Maps* untuk membantu mereka memilih tempat yang tepat. Oleh karena itu, diperlukan analisis sentimen yang mampu mengklasifikasikan ulasan tersebut ke dalam kategori sentimen positif, negatif, atau netral. Hasil dari analisis ini dapat digunakan untuk memberikan rekomendasi *coffee shop* di Daerah Istimewa Yogyakarta. *Support Vector Machine* (SVM) merupakan algoritma yang sangat bergantung pada kualitas fitur yang diberikan. Jika fitur yang digunakan tidak relevan, model *Support Vector Machine* (SVM) tidak akan dapat bekerja dengan baik, namun *Term Frequency-Inverse Document Frequency* (TF-IDF) bisa mengekstrak fitur yang lebih bermakna dari teks, sehingga SVM dapat melakukan klasifikasi dengan lebih akurat.

Penelitian ini menggunakan metode *Support Vector Machine* (SVM) yang dikombinasikan dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Analisis sentiment pada penelitian ini terdiri dari tiga kelas yaitu kelas positif, kelas negatif, dan kelas netral. Tahapan penelitian ini dimulai dengan pengumpulan data ulasan dari *google maps* dengan cara *scraping*, kemudian dilakukan tahapan *preprocessing*. Data hasil dari *preprocessing* selanjutnya diberi label menjadi tiga kelas, yaitu positif, negatif dan netral, selanjutnya tahap ekstraksi fitur dengan TF-IDF, kemudian dilakukan tahap pemodelan dengan metode SVM. Model yang telah dibangun dilakukan pengujian dengan tabel *confusion matrix*. Dari tabel *confusion matrix* tersebut akan diperoleh nilai akurasi, presisi, serta *recall*.

Berdasarkan hasil pengujian dengan menggunakan 1484 data dengan perbandingan data latih dan data uji yaitu 80% : 20%, dimana untuk data latih sebanyak 1.187 data dan data uji sebanyak 297 data, dengan pengujian menggunakan tabel *confusion matrix* dihasilkan nilai akurasi sebesar 85.86%, nilai presisi sebesar 86.24%, serta nilai *recall* sebesar 97.89%.

**Kata kunci** : *Coffee Shop*, Analisis Sentimen, TF-IDF, SVM, *Confusion Matrix*.

## ABSTRACT

*The Special Region of Yogyakarta as one of the largest provinces in Indonesia has many coffee shops that offer a variety of atmospheres, types of coffee, and unique services. This can be seen from the many choices of coffee shops in the Special Region of Yogyakarta that serve coffee with distinctive flavors. Customer reviews play an important role in shaping the image and reputation of a coffee shop, and many people rely on reviews from Google Maps users to help them choose the right place. Therefore, a sentiment analysis is needed that is able to classify these reviews into positive, negative, or neutral sentiment categories. The results of this analysis can be used to provide recommendations for coffee shops in the Special Region of Yogyakarta. Support Vector Machine (SVM) is an algorithm that is highly dependent on the quality of the features provided. If the features used are not relevant, the Support Vector Machine (SVM) model will not be able to work well, but Term Frequency-Inverse Document Frequency (TF-IDF) can extract more meaningful features from the text, so that SVM can classify more accurately.*

*This study uses the Support Vector Machine (SVM) method combined with the Term Frequency-Inverse Document Frequency (TF-IDF) method. Sentiment analysis in this study consists of three classes, namely positive class, negative class, and neutral class. The stages of this study began with collecting review data from Google Maps by scraping, then the preprocessing stage was carried out. The data from the preprocessing results were then labeled into three classes, namely positive, negative and neutral, then the feature extraction stage with TF-IDF, then the modeling stage was carried out with the SVM method. The model that has been built was tested with a confusion matrix table. From the confusion matrix table, the accuracy, precision, and recall values will be obtained.*

*Based on the test results using 1484 data with a comparison of training data and test data of 80%: 20%, where for training data as many as 1,187 data and test data as many as 297 data, testing using the confusion matrix table produced an accuracy value of 85.86%, a precision value of 86.24%, and a recall value of 97.89%.*

**Keywords:** *Coffee Shop, Sentiment Analysis, TF-IDF, SVM, Confusion Matrix.*