

ABSTRAK

Aplikasi *streaming online Disney+ Hotstar* pada *Google Play Store* yang penggunanya sangat banyak membuat perusahaan kesulitan untuk mengkategorikan ulasan ke dalam bentuk sentimen positif dan negatif secara manual sehingga dibutuhkan untuk melakukan pengkategorian tersebut secara efisien. Data ulasan yang diambil sebanyak 8000 data dari ulasan aplikasi *Disney+ Hotstar* pada situs *Google Play Store*, kemudian dataset diberikan label sentimen, data kemudian diolah dengan tahap *preprocessing*, serta dilakukan pembobotan *TF-IDF*. Setelah data berhasil diolah kemudian dilakukan analisis sentimen ulasan dengan menggunakan metode *Support Vector Machine (SVM)* dan *Logistic Regression*. Kedua klasifikasi masing – masing dilakukan pengujian pembagian data dengan menggunakan *KFOLD*, di mana *k* yang digunakan adalah 3,5,7, dan 10. Pada metode *Support Vector Machine* digunakan fungsi *kernel linear*, sedangkan pada metode *Logistic Regression* menggunakan fungsi *sigmoid* untuk menentukan analisis sentimennya. Hasil penelitian didapatkan bahwa metode *SVM* memiliki nilai akurasi, *recall*, *f1-score* terbaik dengan nilai masing-masing yaitu 76.73%, 78.00%, dan 82.74% sedangkan metode *Logistic Regression* memiliki tingkat *precision* terbaik dengan nilai 88.79%.

Kata Kunci: Ulasan, *Disney+ Hotstar*, Analisis Sentimen, *Support Vector Machine*, *kernel linear*, *Logistic Regression*, *sigmoid*

ABSTRACT

Online streaming application Disney+ Hotstar on Google Play Store website which has a lot of users makes it difficult for companies to categorize reviews into positive and negative sentiments manually, so it is necessary to carry out this categorization efficiently. The review data taken was 8000 of reviews data from the application Disney+ Hotstar on Google Play Store website, then the dataset is given a sentiment label, the data then processed in preprocessing stages, and weighting is carried out with TF-IDF. After the data has been successfully processed, then the classification of sentiment analysis of the review is carried out using the method Support Vector Machine (SVM) and Logistic Regression. The two classifications were each tested using data splitting method called KFOLD, where k used are 3,5,7, and 10. Linear kernel is used in Support Vector Machine classification method, while sigmoid function used in Logistic Regression method to determine the sentiment analysis. The research results showed that the SVM method has value of accuracy, recall, f1-score best with respective values namely 76.73%, 78.00%, and 82.74% while the Logistic Regression has the best precision value with a score of 88.79%.

Keywords: Review, Disney+ Hotstar, Sentiment Analysis, Support Vector Machine, kernel linear, Logistic Regression, sigmoid