

## INTISARI

### **KLASIFIKASI *REVIEW SOFTWARE* PADA *GOOGLE PLAY* MENGGUNAKAN PENDEKATAN ANALISIS SENTIMEN**

Oleh

Firdauz Fanani

13/347109/PA/15160

*Review* pada *google play* yang merupakan salah satu fitur yang digunakan untuk memberikan suatu penilaian terhadap suatu aplikasi. Namun terdapat penyalahgunaan oleh pihak-pihak tertentu untuk mengambil keuntungan atau merugikan pihak lain dengan membuat *review SPAM*. *Review SPAM* dapat merugikan para pengguna dan mengganggu para *developer* yang menggunakan *review* tersebut sebagai bahan dalam melakukan evaluasi terhadap aplikasinya. Untuk itu dibutuhkan sistem yang dapat mengklasifikasikan *review-review* tersebut sebagai *review SPAM* atau bukan *SPAM(HAM)*.

Dalam penelitian ini diterapkan tahap-tahap *preprocessing* yang akan mengolah data-data mentah dari suatu *review* menjadi data yang siap diolah dan juga dibuat sebuah model yang merepresentasikan data-data yang sudah dilakukan *preprocessing* menjadi matriks lalu dilakukan pembuatan vektor kata-kata pada *review* tersebut. Algoritma yang digunakan dalam klasifikasi ini adalah algoritma *Support Vector Machine(SVM)* dan *Naïve Bayes Classifier(NBC)*. Pengujian akhir sistem dilakukan dengan menguji dan membandingkan performa dari algoritma *Support Vector Machine* dan *Naïve Bayes Classifier*.

Hasil akhir dari penelitian ini menunjukkan performa sistem klasifikasi dengan algoritma *Support Vector Machine* menghasilkan akurasi lebih baik dibandingkan dengan menggunakan algoritma *Naïve Bayes Classifier*.

Kata kunci: Klasifikasi *Review*, *HAM*, *SPAM*, *Preprocessing*, *Support Vector Machine(SVM)*, *Naïve Bayes Classifier(NBC)*.

## **ABSTRACT**

### **REVIEW SOFTWARE CLASSIFICATION ON GOOGLE PLAY USING SENTIMENT ANALYSIS APPROACH**

By

Firdauz Fanani  
13/347109/PA/15160

Review on google play which is one of the features used to provide an assessment of an application. However there is a misuse by certain parties to take advantage or harm others by making SPAM reviews. SPAM reviews may harm users and disrupt developers who use that reviews as a material in evaluating their applications. This requires a system that can classify these reviews as SPAM or non-SPAM (HAM) reviews.

In this study applied preprocessing stages that will process raw data from reviews into ready-to-process data and also created a model that represents data that has been done preprocessing into a matrix and then make vector of words on the reviews. The algorithms used in this classification are the Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC) algorithms. The final test of the system is done by testing and comparing the performance of the Support Vector Machine and Naïve Bayes Classifier algorithms.

The final result of this research shows the performance of classification system with Support Vector Machine algorithm yield accuracy better than using Naïve Bayes Classifier algorithm.

**Keywords:** Review Classification, HAM, SPAM, Preprocessing, Support Vector Machine (SVM), Naïve Bayes Classifier (NBC).