

ABSTRAK

Nama : Nove Suarti Hia
Program Studi : Informatika
Judul : Implementasi Algoritma *Random Forest* pada Klasifikasi *Review* Aplikasi di *Google Play Store*
Dosen Pembimbing : M. Soleh, M.Kom

Google play Store adalah sebuah sistem yang terdiri dari aplikasi, buku, lagu, dan lain-lain. *Google Play* dilengkapi dengan adanya fitur berisi *review* dari para pengguna yang dapat digunakan untuk melihat *review* dari pengguna aplikasi. *Review* sangat sering digunakan sebagai alat yang efektif dan efisien dalam menemukan informasi terhadap suatu produk atau jasa. Dengan banyaknya *review* yang ada di situs *google play store* mengakibatkan pihak perusahaan sulit untuk memperoleh informasi semua *review*, karena membutuhkan waktu yang lama jika membaca satu per satu *review*. Oleh sebab itu, diperlukan sebuah metode atau teknik khusus yang mampu mengklasifikasikan *review* tersebut secara otomatis. Pada penelitian ini dilakukan proses klasifikasi dengan menggunakan algoritma *Random Forest* yang merupakan salah satu metode dari *machine learning*. Metodologi yang digunakan untuk melakukan klasifikasi dimulai dari *data collecting*, *preprocessing*, *feature selection*, klasifikasi dan pengukuran akurasi. Proses klasifikasi pada *random forest* berawal dari memecah data sampel yang ada kedalam *decision tree* secara acak. Sebelum proses klasifikasi, ada beberapa tahapan yaitu *Preprocessing* yang terdiri dari *case folding*, *cleaning*, *filtering*, tokenisasi, *stemming* dan penentuan atribut. Tahapan selanjutnya yaitu melakukan *modelling* dengan algoritma *Random forest* yang menggunakan *Decision Tree* untuk melakukan proses seleksi. Data pada penelitian ini dibagi menjadi dua yaitu, 80% untuk data *training* dan 20% untuk data *testing*. Penelitian ini menghasilkan akurasi prediksi sebesar 89%.

Kata kunci : *Decision Tree*, Klasifikasi, *Machine Learning*, *Preprocessing*, *Random Forest*.

ABSTRACT

The Google Play Store is a system consisting of applications, books, songs, and more. Google Play is equipped with a feature that contains reviews from users that can be used to view reviews from application users. Reviews are very often used as an effective and efficient tool in finding information on a product or service. With so many reviews on the Google Play Store site, it is difficult for the company to get information on all reviews, because it takes a long time to read one review at a time. Therefore, we need a special method or technique that is able to classify these reviews automatically. In this study, the classification process was carried out using the Random Forest algorithm, which is one of the methods of machine learning. The methodology used to classify starts from data

collecting, preprocessing, feature selection, classification and measurement accuracy. The classification process in random forest begins with breaking down the existing sample data into a random decision tree. Before the classification process, there are several stages, namely Preprocessing which consists of case folding, cleaning, filtering, tokenization, stemming and attribute determination. The next step is modeling the Random forest algorithm using the Decision Tree to carry out the selection process. The data in this study were divided into two, namely, 80% for training data and 20% for testing data. This study resulted in a predictive accuracy of 89%.

Keywords : Classification, Decision Tree, Machine Learning, Preprocessing, Random Forest.

