

Klasifikasi ulasan pengguna aplikasi mobile berbahasa Indonesia: Studi kasus PT PQR = Indonesian mobile application reviews classification: Case study PT PQR

Yudo Ekanata, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20468283&lokasi=lokal>

Abstrak

ABSTRAK

Jumlah ulasan pengguna untuk sebuah aplikasi mobile bisa mencapai ribuan sehingga membutuhkan waktu banyak bagi pengembang aplikasi untuk memilah-milah dan mencari informasi yang penting untuk pengembangan aplikasi selanjutnya. Oleh karena itu penelitian ini bertujuan untuk melakukan klasifikasi secara otomatis terhadap ulasan pengguna aplikasi mobile yang dimiliki oleh PT PQR. Klasifikasi otomatis yang dilakukan pada penelitian ini menggunakan pendekatan machine learning. Fitur-fitur yang diekstrak dari data adalah unigram, bigram, star rating, panjang ulasan, serta rasio jumlah kata bersentimen positif dan negatif. Untuk algoritma klasifikasi digunakan Na ve Bayes, Support Vector Machine, Logistic Regression dan Decision Tree. Hasil percobaan menunjukkan bahwa fitur bigram dan star rating berdampak negatif pada performa klasifikasi. Algoritma Na ve Bayes menghasilkan precision terbaik sebesar 85,3, sedangkan algoritma Logistic Regression menghasilkan recall dan F-measure terbaik sebesar 85.

ABSTRACT

The number of user reviews for a mobile app can reach thousands so it will take a lot of time for app developers to sort through and find information that is important for further app development. Therefore, this study aims to classify automatically mobile application user reviews of PT PQR. Automatic classification conducted in this study using machine learning approach. The features extracted from the data are unigram, bigram, rating score, review length, as well as the ratio of the number of words contain positive and negative sentiment. For classification algorithms, we used Na ve Bayes, Support Vector Machine, Logistic Regression and Decision Tree. The experimental results show that bigram and star rating has negative impact on the classifier performance. Na ve Bayes algorithm gives the best precision of 85,3 , while Logistic Regression gives the best recall and F measure of 85.