

COMPARISON OF SUPPORT VECTOR MACHINE AND NAÏVE BAYES ALGORITHMS IN ANALYZING USER SENTIMENTS OF THE HAJI PINTAR APPLICATION

ABSTARCT

This study discusses the performance of two algorithms in analyzing the sentiment of Hajj Pintar application users, so that comparative analysis is needed to determine the algorithm that has optimal performance in sentiment analysis. The research method applied is the comparative experimental method. The purpose of the study was to compare the algorithms of Support Vector Machine and Naïve Bayes in analyzing the sentiment of users of the Haji Pintar application. The dataset was obtained from the Google Play Store with a web scraping technique of 2091 data. The evaluation was carried out based on evaluation metrics, namely accuracy, precision, recall, and F1-Score with the proportion of test data and training data with ratios of 10:90, 20:80 and 30:70. The test results with test data and training data with a ratio of 10:90, the SVM algorithm obtained an accuracy score of 90.48%, precision of 88.64%, recall of 92.86%, and F1-Score of 90.70%. Meanwhile, the NB algorithm only obtained an accuracy score of 90.48%, precision of 86.96%, recall of 95.24%, and F1-Score of 90.91%. The test results with test data and training data with a ratio of 20:80, the SVM algorithm obtained an accuracy score of 91.37%, precision of 90.17%, recall of 92.86%, and F1-Score of 91.50%. Meanwhile, the NB algorithm only produced an accuracy value of 90.18%, precision of 87.29%, recall of 94.05%, and F1-Score of 90.54%. The test results with test data and training data with a ratio of 30:70, the SVM algorithm produced an accuracy value of 91.47%, precision of 89.73%, recall of 93.65%, and F1-Score of 91.65%, while for the NB algorithm only obtained an accuracy value of 89.48%, precision of 85.66%, recall of 94.84%, and F1-Score of 90.02%. Based on the results of this algorithm comparison, the performance level of the SVM algorithm is proven to be more optimal than the NB algorithm in analyzing the sentiment of users of the Hajj Pintar application, in each division of test data and training data with a ratio of 10:90, 20:80 and 30:70.

Keywords: Sentiment Analysis, Haji Pintar application, Comparison, Support Vector Machine, Naïve Bayes.

PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES* DALAM MENGANALISIS SENTIMEN PENGGUNA APLIKASI HAJI PINTAR

ABSTRAK

Penelitian ini membahas kinerja dari dua algoritma dalam menganalisis sentimen pengguna aplikasi Haji Pintar, sehingga diperlukan analisis komparatif untuk menentukan algoritma yang memiliki kinerja yang optimal dalam analisis sentimen. Metode penelitian yang diterapkan adalah metode eksperimen komparatif. Tujuan penelitian untuk melakukan perbandingan algoritma *Support Vector Machine* dan *Naïve Bayes* dalam menganalisis sentimen pengguna aplikasi Haji Pintar. Dataset diperoleh dari *Google Play Store* dengan teknik *web scraping* sebanyak 2091 data. Evaluasi dilakukan berdasarkan metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-Score* dengan proporsi data uji dan data latih dengan rasio 10:90, 20:80 dan 30:70. Hasil pengujian dengan data uji dan data latih dengan rasio 10:90, algoritma SVM memperoleh nilai *accuracy* 90.48%, *precision* 88.64%, *recall* 92.86%, dan *F1-Score* 90.70%. Sedangkan, algoritma NB hanya memperoleh nilai *accuracy* 90.48%, *precision* 86.96%, *recall* 95.24%, dan *F1-Score* 90.91%. Hasil pengujian dengan data uji dan data latih dengan rasio 20:80, algoritma SVM memperoleh nilai *accuracy* 91.37%, *precision* 90.17%, *recall* 92.86%, dan *F1-Score* 91.50%. Sedangkan, algoritma NB hanya menghasilkan nilai *accuracy* 90.18%, *precision* 87.29%, *recall* 94.05%, dan *F1-Score* 90.54%. Hasil pengujian dengan data uji dan data latih dengan rasio 30:70, algoritma SVM menghasilkan nilai *accuracy* 91.47%, *precision* 89.73%, *recall* 93.65%, dan *F1-Score* 91.65%, sedangkan untuk algoritma NB hanya memperoleh nilai *accuracy* 89.48%, *precision* 85.66%, *recall* 94.84%, dan *F1-Score* 90.02%. Berdasarkan hasil komparasi algoritma ini, tingkat kinerja algoritma SVM terbukti lebih optimal dibandingkan algoritma NB dalam menganalisis sentimen pengguna aplikasi Haji Pintar, pada setiap pembagian data uji dan data latih dengan rasio 10:90, 20:80 serta 30:70.

Kata Kunci: Analisis Sentimen, Aplikasi Haji Pintar, Perbandingan, *Support Vector Machine*, *Naïve Bayes*