

ANALISIS EMOSI *TWITTER* MENGGUNAKAN METODE SELEKSI FITUR *MUTUAL INFORMATION* DAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM)

Oleh :

Ahmad Marzuqi Yasykur Luthfi
09021281823092

ABSTRAK

Emosi adalah hal yang bersifat umum dan penting dalam kehidupan yang membentuk perilaku manusia. Emosi seseorang bisa diidentifikasi menggunakan data tekstual yaitu berupa teks, karena teks dapat digunakan untuk berkomunikasi dan menyampaikan informasi. Seleksi fitur digunakan untuk mengurangi dimensi ruang fitur dengan cara melakukan pembobotan pada semua fitur kemudian fitur dengan bobot yang kurang dari ambang batas akan dieliminasi. Hal ini bertujuan untuk meningkatkan akurasi serta efisiensi waktu komputasi pada proses klasifikasi teks. Pada penelitian ini, metode Mutual Information digunakan pada proses klasifikasi teks berupa sentimen emosi pada Twitter menggunakan algoritma Support Vector Machine (SVM). Kemudian, akan dilakukan analisis perbandingan pada setiap model klasifikasi berdasarkan hasil evaluasi yang didapat. Hasil penelitian menunjukkan penggunaan metode seleksi fitur mampu memberikan peningkatan akurasi serta mengurangi waktu komputasi. Penggunaan metode seleksi fitur Mutual Information pada algoritma SVM dengan kernel linear menghasilkan kinerja terbaik dengan rata-rata accuracy 0.71, precision 0.72, recall 0.71, f-measure 0.71 dan waktu komputasi 1.16 detik.

Kata Kunci : Klasifikasi Teks, Seleksi Fitur, *Support Vector Machine* (SVM)

Pembimbing I

Indralaya, 14 Juli 2023
Pembimbing II

Samsuryadi, M.Kom., Ph.D.
NIP. 197102041997021003

Rizki Kurniati, M.T.
NIP. 199107122019032016

Mengetahui,
Ketua Jurusan Teknik Informatika

Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

TWITTER EMOTION ANALYSIS USING MUTUAL INFORMATION FEATURE SELECTION METHOD AND SUPPORT VECTOR MACHINE (SVM) ALGORITHM

By :

Ahmad Marzuqi Yasykur Luthfi
09021381823092

ABSTRACT

Most texts have a large number of features. However, the features contained in the text mostly have a low level of relevance and even contain noise which can later reduce the accuracy of the results. Feature selection is used to reduce the dimensions of feature space by weighting all features then features with lower weights than threshold will be eliminated. It aims to improve the accuracy and efficiency of computational time in the text classification process. In this research, selection method Mutual Information were used in the text classification process in the form of Twitter emotion sentiment using the Support Vector Machine (SVM) algorithm. Then, a comparative analysis will be carried out on each classification model based on the evaluation results obtained. The results showed that the use of the feature selection method was able to increase accuracy and reduce computation time. The use of the Mutual Information feature selection method on the SVM algorithm with a linear kernel give the best performance with average of accuracy 0.71, precision 0.72, recall 0.71, f-measure 0.71 and computation time 1.17 seconds.

Key Word : Text Classification, Feature Selection, Support Vector Machine (SVM)

Advisor I

Indralaya, 24 June 2023
Advisor II

Samsuryadi, M.Kom., Ph.D.
NIP. 197102041997021003

Rizki Kurniati, M.T.
NIP. 199107122019032016

Approved,
Head of Informatics Engineering Department

Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003