Abstract
Lyrics are the most important component of a song that can be analyzed as data consisting of thoughts, moods, emotions or feelings related to a particular topic. The lyrics also provide a wider display of songs than simpler forms of data such as artists, years, titles, genres and albums. Determination of emotions in the lyrics contains words that correspond to emotional characteristics. One of the lyrics can be categorized into happy and sad emotions based on the words contained in the lyrics. Many studies have classified song emotions based on lyrics. However, these studies have poor accuracy and are still not optimal. This study provides a solution to this problem by classifying emotional lyrics based on lyrics using the support vector machine with the gain ratio feature selection. Classification using SVM with linear kernel provides a better value of accuracy on the classification of lyric text data. The first step carried out in the lyric text data is preprocessing. Then do the weighting of words and feature selection using the gain ratio. The results of this stage are used as input for the SVM classification. This study obtained accuracy with the use of preprocessing, weighting words using tf-idf, feature-based selection gain ratio, SVM classification and k-fold 10 which amounted to 70.23%.

Keywords : emotion classification, lyrics, tf-idf, gain ratio, support vector machine