PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi emosi berbasis suara manusia dengan menggunakan *Mel Frequency Cepstral Coefficients* (MFCC), *Delta*, dan *Delta-Delta* sebagai metode utama ekstraksi fitur, serta algoritma *Support Vector Machine* (SVM) untuk proses klasifikasi. Permasalahan utama yang diangkat adalah bagaimana kinerja ketiga jenis fitur tersebut dalam merepresentasikan emosi suara, serta efektivitas kernel yang digunakan dalam SVM, khususnya perbandingan antara kernel *Polynomial* dan *Radial Basis Function* (RBF).Dataset yang digunakan adalah *Ryerson Audio-Visual Database of Emotional Speech and Song* (RAVDESS), yang mencakup delapan kelas emosi: *neutral, calm, happy, sad, angry, fearful, disgust,* dan *surprised*. Proses dimulai dengan *pre-processing* (normalisasi dan penghilangan noise), diikuti ekstraksi fitur MFCC beserta turunannya (*Delta* dan *Delta-Delta*). Model SVM kemudian diuji menggunakan berbagai kombinasi parameter dan kernel untuk mengevaluasi performanya.Hasil penelitian menunjukkan bahwa model terbaik adalah SVM dengan kernel RBF, parameter C=10 dan gamma=0.1, tanpa *class weight*, serta menggunakan fitur lengkap (MFCC, *Delta, Delta-Delta* dengan satu nilai *mean* per audio), yang menghasilkan akurasi tertinggi sebesar 84,33% dan performa seimbang di seluruh kelas emosi. Kernel RBF terbukti unggul karena kemampuannya memodelkan hubungan *non-linear* dalam data audio. Selain itu, kombinasi fitur lengkap lebih efektif dalam menangkap karakteristik spektral dan temporal suara dibandingkan penggunaan MFCC saja. Temuan ini

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

mengindikasikan bahwa fitur turunan (*Delta* dan *Delta-Delta*) memberikan kontribusi signifikan dalam meningkatkan akurasi klasifikasi emosi berbasis suara.

Kata Kunci: *Klasifikasi Emosi, MFCC, Delta, Delta-Delta, SVM, Kernel RBF, RAVDESS, Ekstraksi Fitur, Audio Pre-processing.*

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRACT

This study aims to develop a human speech-based emotion classification system using Mel Frequency Cepstral Coefficients (MFCC), Delta, and Delta-Delta as the main feature extraction methods, along with the Support Vector Machine (SVM) algorithm for the classification process. The primary research question is how well these three types of features represent vocal emotions, and the effectiveness of different SVM kernels, particularly the comparison between the Polynomial and Radial Basis Function (RBF) kernels. The dataset used is the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS), which includes eight emotion classes: neutral, calm, happy, sad, angry, fearful, disgust, and surprised. The process begins with pre-processing (normalization and noise reduction), followed by feature extraction using MFCC and its derivatives (Delta and Delta-Delta). The SVM model was then evaluated using various combinations of parameters and kernels to assess its performance. The results show that the best-performing model is the SVM with an RBF kernel, using parameters C=10 and gamma='scale', without class weighting, and employing the full feature set (MFCC, Delta, Delta-Delta, using a single mean value per audio). This configuration achieved the highest accuracy of 84.33%, with balanced performance across all emotion classes. The RBF kernel proved to be superior due to its ability to model non-linear relationships in audio data. Additionally, the complete feature set was more effective in capturing the spectral and temporal characteristics of speech compared to using MFCC alone. These findings suggest that derivative features (Delta and Delta-Delta) contribute significantly to improving the accuracy of speech-based emotion classification.

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

Keywords: *Emotion Classification, MFCC, Delta, Delta-Delta, SVM, RBF Kernel, RAVDESS, Feature Extraction, Audio Pre-processing.*

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI