ABSTRAK

Deteksi depresi secara otomatis melalui analisis ekspresi wajah dan gerakan kepala menjadi bidang yang semakin penting dalam pengembangan sistem cerdas berbasis kecerdasan buatan. Penelitian ini bertujuan untuk mengoptimalkan teknik pembagian data (*data splitting*) pada algoritma *Random Forest* dalam upaya meningkatkan akurasi dan efisiensi klasifikasi episode depresi. Data yang digunakan berasal dari *dataset Beyond Smile Challenge* yang terdiri dari lebih dari 286.074 entri data gerak wajah dan kepala dari 25 *partisipan*. Proses penelitian melibatkan tahapan pra-pemrosesan data, ekstraksi fitur penting seperti *Action Units* (*AU*), struktur wajah, dan probabilitas mata terbuka, serta penerapan teknik normalisasi. Evaluasi performa model dilakukan menggunakan pendekatan *Leave-One-Participant-Out* (*LOPO*) *Cross-Validation* untuk menguji generalisasi antar individu. Hasil evaluasi menunjukkan bahwa model *Random Forest* mampu mencapai akurasi rata-rata sebesar 84,1%, dengan nilai precision 74,7%, recall 86,2%, F1-score 78,5%, dan AUC 90,3%. Temuan ini membuktikan bahwa strategi *splitting data* yang tepat dapat secara signifikan meningkatkan kinerja algoritma *Random Forest* dalam klasifikasi kondisi depresi berbasis data *non-verbal*. Penelitian ini memberikan kontribusi dalam pengembangan sistem deteksi depresi yang andal serta menawarkan panduan optimalisasi *splitting data* untuk studi lanjutan.

Kata Kunci: *Random Forest, Splitting Data,* Ekspresi Wajah, Gerak Kepala, Deteksi Depresi, *LOPO Cross-Validation*.

ABSTRACT

Depression detection through automated analysis of facial expressions and head movements has become an increasingly important field in the development of AI-based intelligent systems. This study aims to optimize data splitting techniques in the Random Forest algorithm to improve the accuracy and efficiency of depression episode classification. The data used was sourced from the Beyond Smile Challenge dataset, which consists of over 280,000 entries of facial and head movement data from 25 participants. The research process involves data pre-processing, extraction of key features such as Action Units (AU), facial structure, and eye-opening probability, as well as the application of normalization techniques. Model performance was evaluated using Leave-One-Participant-Out (LOPO) Cross-Validation to assess generalization across individuals. The results show that the Random Forest model achieved an average accuracy of 84.1%, with a precision of 74.7%, recall of 86.2%, F1-score of 78.5%, and AUC of 90.3%. These findings demonstrate that an appropriate data splitting strategy can significantly enhance the performance of the Random Forest algorithm in non-verbal depression classification. This research contributes to the development of reliable depression detection systems and provides guideline for optimizing data splitting in future studies.

Keywords: *Random Forest, Data Splitting, Facial Expressions, Head Movements, Depression Detection, LOPO Cross-Validation.*