Abstract

The discussion has seen that emotion recognition poses a significant challenge to human-computer interaction and mental health applications, and genuine tools are required in recognition of emotions in real-time to facilitate users' experience and assist in psychological appraisal. Taking into consideration the issue, the researcher has been employing machine learning methods to distinguish the emotions that are complicated states affecting the facial expression, voice tones, and the circumstances. The FER-2013 dataset of gray-scale images of emotional faces (anger, disgust, fear, happiness, sadness, surprise, and neutral) is also collected using the secondary method of data collection. The key agenda is to establish the effectiveness of the supervised learning models in the classification of emotions to be used in affective computing. Models that the researcher has been employing include the logistic regression, random forest, and gradient boosting models, which were created using Python and scikit-learn.

The researcher adheres to a logical approach in this study, including such data processing, exploratory data analysis (EDA), model development, evaluation, and validation. The data had been curated with missing values, image normalization, and stratified subdivision in order to overcome the imbalance among the classes. It was found through EDA that these image characteristics, such as pixel intensity distributions in the areas of the face (such as the eyes, mouth, etc.), are highly predictive of emotions, which is in line with psychological theories. The models were written in Python using libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn, and unit tested using Pytest. Performance was measured using accuracy, precision, recall, F1-score, and confusion matrices, and generalizability was obtained by using k-fold cross-validation.

The critical review of findings reveals that the Gradient Boosting model has had cohesive operation as indicated by 65 percent accuracy on the test set, which is better than the performance of the Random Forest and the Logistic Regression models. Such is observed in the model testing, whereby the Logistic Regression is more consistent in cross-validation (~0.62 mean accuracy), meaning it has a high level of reliability, albeit lower than the peak. These findings indicate the compromise existing between complex ensemble models, which represent subtle patterns in the faces, and simpler models, which are more understandable in real-time applications.

The unit testing process has assisted in ensuring that the pipeline created has been developed in the correct way, and end-to-end workflow validation confirms that the pipeline is correct regarding emotion predictions. One may say that this project increased technical capabilities in the machine learning process, image information analysis in affective computing, and performance assessment. The largest problems were the control of the imbalance in classes, optimization of hyperparameters, and understanding the model. ML pipeline implementation through modularity, which promoted the efficiency and reproducibility of the code, and comparison of models provided cogent information on both accuracy and usability in interactive systems. The researcher will utilize the multimodal datasets (e.g., audio and video) in the future to increase the accuracy of recognition.