SERAB Phase 2 Submission Report

Introduction

Our approach integrates feature engineering, ensemble modeling, hyperparameter optimization, enhanced evaluation metrics, and model persistence to boost model performance and efficiency.

1 Feature Engineering (PCA)

Original Dimensions: Embeddings in the training set had an initial shape of $(5144, 1024)$.
Reduced Dimensions: After applying Principal Component Analysis (PCA), the dimensions were reduced to $(5882, 518)$. PCA preserved 95% of the variance, focusing on essential features.

Result: PCA enabled efficient handling of high-dimensional audio data, reducing computational requirements without significant loss of variance.

2 Ensemble Modeling (Stacking and Voting)

Stacking Classifier Accuracy: 70%
Voting Classifier Accuracy: 68%

Result: Stacking achieved a notable improvement over individual classifiers, demonstrating the power of combining multiple models for higher accuracy in audio classification.

3 Hyperparameter Optimization

Using `RandomizedSearchCV`, we efficiently optimized hyperparameters, running 27 out of the initially set 98 iterations due to a limited parameter space.

Best Model Performance

- **Test Accuracy:** 68%
- **Precision:** 68%
- **Recall:** 68%
- **F1 Score:** 68%

Result: Balanced performance across multiple metrics suggests that hyperparameter tuning provided a stable, well-optimized model.

4 Enhanced Evaluation Metrics

Evaluation on CREMA-D Dataset:

- **Test Accuracy:** 70.2%
- **Unweighted Average Recall (UAR):** 70.4%

Result: The model generalized well to unseen data, as evidenced by balanced accuracy and recall. This result indicates its effectiveness in real-world audio classification.

5 Model Persistence

The trained model was saved and successfully reloaded using `joblib`, ensuring easy deployment and future usage.

6 Final Results on CREMA-D Dataset

- **Test Accuracy:** 70.2%
- **Test UAR:** 70.4%

7 Impact on Execution Time and Accuracy

Execution Time: PCA's dimensionality reduction decreased overall execution time by simplifying high-dimensional data, particularly improving training and testing efficiency.

Accuracy: Ensemble models (Stacking and Voting) increased accuracy over baseline classifiers. The stacking classifier's 70% accuracy outperformed individual models.

Overall Outcome: The updated code achieved competitive accuracy, particularly with the stacking ensemble, and reduced execution time due to PCA and optimized hyperparameters. Together, these improvements streamlined the model and enhanced performance.

Conclusion

Our updated approach in SERAB Phase 2 improved both model accuracy and efficiency, confirming the effectiveness of advanced feature engineering, ensemble methods, and optimization techniques for audio classification. This report highlights the potential of these methods for further applications and advancements in the field.
