SPOKEN GRAMMAR SCORING SYSTEM - MODEL REPORT 1

Execute top 10 features based on importance

NAME: UDAYAN SAHA

COLLEGE: KAZI NAZRUL UNIVERSITY [A state govt, university, Asansol, West Bengal]

CURRENT STATUS: 4TH YEAR B. TECH COMPUTER SCIENCE STUDENT

SPECIALIZATION: DATA SCIENCE

Abstract

This report documents the development and evaluation of a machine learning model designed to score the grammatical quality of spoken English audio samples on a scale of 0 to 5. The model processes engineered audio features and predicts grammar quality using a tuned ***Gradient Boosting Regressor***.

1. Introduction

The ability to assess spoken grammar automatically has significant implications for educational and recruitment platforms. In this work, we leverage acoustic features extracted from .wav files and apply regression models to predict human-assigned grammar scores.

2. Dataset and Preprocessing

Audio features were extracted from 3,000+ .wav files using the opensmile toolkit. Post-processing included handling nulls, filtering low-variance features, and standardization. The target label was the grammar score between 0 and 5.

Train-test split was done using an 80-20 ratio.

3. Feature Engineering

Initial feature space had 6,373 dimensions. After cleaning and low-variance filtering, it was reduced to 455 dimensions. StandardScaler was used to normalize features.

4. Model and Hyperparameter Tuning

We selected ***Gradient Boosting Regressor*** and optimized its parameters using ***RandomizedSearchCV*** with the following grid:

```
param_grid = {
    'n_estimators': [100, 200, 300],
    'learning_rate': [0.01, 0.05, 0.1],
    'max_depth': [3, 5, 7],
    'min_samples_split': [2, 5, 10],
    'subsample': [0.6, 0.8, 1.0]
}
```

5. Evaluation

The model was evaluated using:

- **R^2 Score**: 0.2781
- **Mean Squared Error (MSE)**: 0.8679

These metrics indicate moderate model performance, suitable for rank-based evaluation rather than precise prediction.

6. Top 10 Features

The most informative features contributing to the grammar score prediction:

Rank	Feature	Importance
1	F0semitoneFrom27.5Hz_sma3_amean	0.0349
2	pcm_RMSenergy_sma_de	0.0262
3	F0env_sma3_amean	0.0245
4	pcm_fftMag_mfcc_sma[1]_linregc1	0.0223
5	pcm_fftMag_mfcc_sma[2]_linregc1	0.0201
6	pcm_fftMag_mfcc_sma[3]_linregc1	0.0196
7	pcm_fftMag_mfcc_sma[4]_linregc1	0.0178
8	pcm_fftMag_mfcc_sma[5]_linregc1	0.0164
9	pcm_fftMag_mfcc_sma[6]_linregc1	0.0153
10	pcm_fftMag_mfcc_sma[7]_linregc1	0.0139

7. Conclusion

The tuned Gradient Boosting Regressor demonstrates promise in estimating grammar quality from speech. Further performance improvements could be achieved via ensemble modelling, larger training data, and leveraging deep acoustic embeddings like wav2vec.