**Report on Accuracy Improvement Journey**

**Project Overview:**

The project focuses on detecting and predicting gerrymandering using machine learning models. The primary goal is to improve the accuracy of the predictions step by step, employing various techniques and model enhancements.

**Initial Model Setup:**

1. **Initial Model: RandomForestClassifier**
   * **Features Used:** margin16, margin18, vote\_to\_seat\_ratio
   * **Initial Accuracy:** ~53.44%
   * **Observations:** The model struggled with classifying the "high" and "low" gerrymandering categories. The accuracy was moderate, but there was significant room for improvement.

**Step 1: Hyperparameter Tuning with RandomizedSearchCV**

* **Approach:**
  + Performed hyperparameter tuning using RandomizedSearchCV on the RandomForestClassifier.
  + Optimized parameters like n\_estimators, max\_depth, min\_samples\_split, etc.
* **Results:**
  + **Improved Accuracy:** ~53.44% (same as initial due to random nature)
  + **Observations:** Although the accuracy didn't improve significantly, this step helped in fine-tuning the model's parameters.

**Step 2: Implementing a Stacking Classifier**

* **Approach:**
  + Implemented a StackingClassifier with base models: RandomForestClassifier, XGBClassifier, and LGBMClassifier.
  + The final estimator was set as RandomForestClassifier.
* **Results:**
  + **Stacking Classifier Accuracy:** ~58.78%
  + **Observations:** The stacking model showed a moderate improvement in accuracy. The "mid" class showed better performance, while the "high" and "low" classes still needed improvement.

**Step 3: Feature Engineering with Polynomial Features**

* **Approach:**
  + Generated polynomial and interaction features using PolynomialFeatures (degree=2).
  + Retrained the stacking classifier with these new features.
* **Results:**
  + **Accuracy with Polynomial Features:** ~61.83%
  + **Observations:** This step brought a significant improvement, especially in the "mid" class. However, the "high" and "low" classes remained challenging to classify accurately.

**Conclusion:**

Through systematic improvements—hyperparameter tuning, stacking classifiers, and introducing polynomial features—the accuracy of the model increased from an initial ~53.44% to ~61.83%. Each step contributed to the overall enhancement, with polynomial features playing a crucial role in the final improvement. Further work may focus on addressing the performance in the "high" and "low" categories, possibly through advanced feature engineering or model ensembling techniques.