# Model Performance Documentation

## 1. Model Overview

- Model Type: Gradient Boosting Regressor  
- Training Dataset: Original dataset with specified features  
- Target Variable: Case\_Disposal\_Rate  
- Features Used:  
 - Pending\_to\_Disposed\_Ratio  
 - Normalized\_Pending\_Case\_Growth  
 - Pending\_Case\_Growth  
 - Normalized\_Pending\_to\_Disposed\_Ratio  
 - Court\_Efficiency\_Index  
 - Pending cases for a period of 0 to 1 Years  
 - Normalized\_Old\_Pending\_Percentage  
 - Normalized\_Delay\_Ratio

## 2. Preprocessing Steps

1. Imputation: Used SimpleImputer with mean strategy to handle any missing values in the numeric feature columns.  
2. Data Splitting: Split the dataset into an 80% training set and a 20% test set.  
3. Random State: Set random\_state=42 for reproducibility.

## 3. Model Hyperparameters

- n\_estimators: 75  
- learning\_rate: 0.05  
- max\_depth: 3  
- min\_samples\_split: 20  
- min\_samples\_leaf: 10  
- subsample: 0.7  
- random\_state: 42

## 4. Performance on Holdout Test Set

Evaluated on the 20% holdout test set.  
  
- Mean Squared Error (MSE): 0.0002  
- Mean Absolute Error (MAE): 0.0032  
- R-squared (R²): 0.9726  
  
These metrics indicate excellent performance on the test set, with the model explaining approximately 97.26% of the variance in the target variable.

## 5. Cross-Validation Results

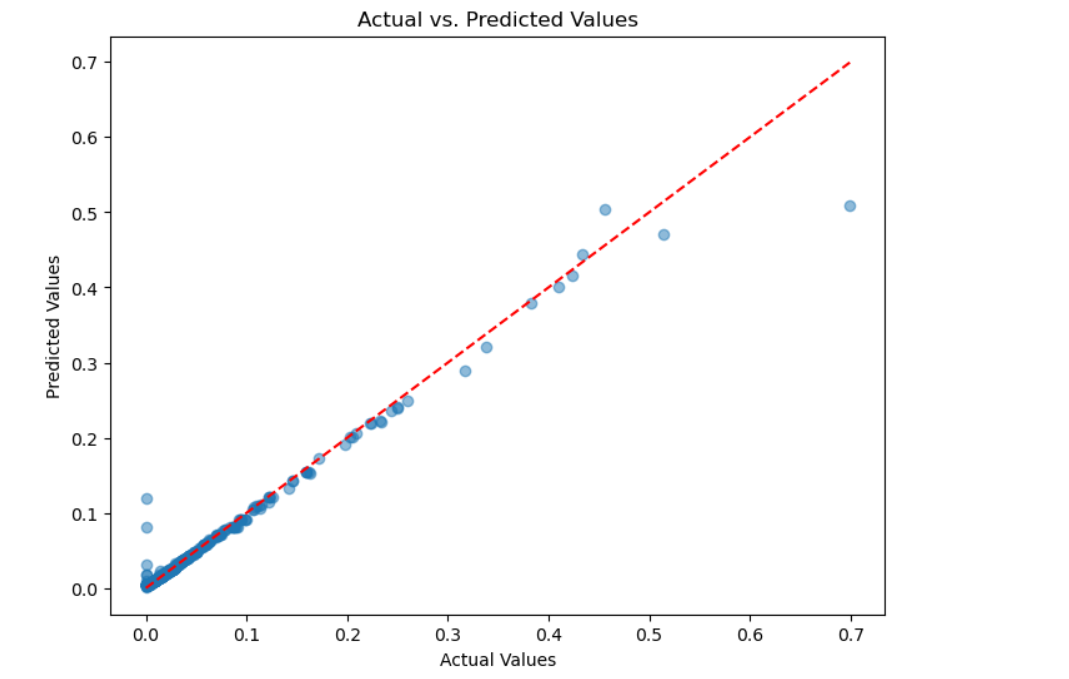
- Cross-validated R² scores: [0.7611, 0.9985, 0.9865, 0.8302, 0.9912, 0.9830, 0.9867, 0.9967, 0.9711, 0.9969]  
- Mean R²: 0.9502  
- Standard Deviation of R²: 0.0792

## 6. Bootstrapping Results

- Bootstrapped R² scores: [0.9902, 0.9905, 0.9937, 0.9862, 0.9851, 0.9888, 0.9899, 0.9842, 0.9843, 0.9809]  
- Mean Bootstrapped R²: 0.9874  
- Standard Deviation of Bootstrapped R²: 0.0037

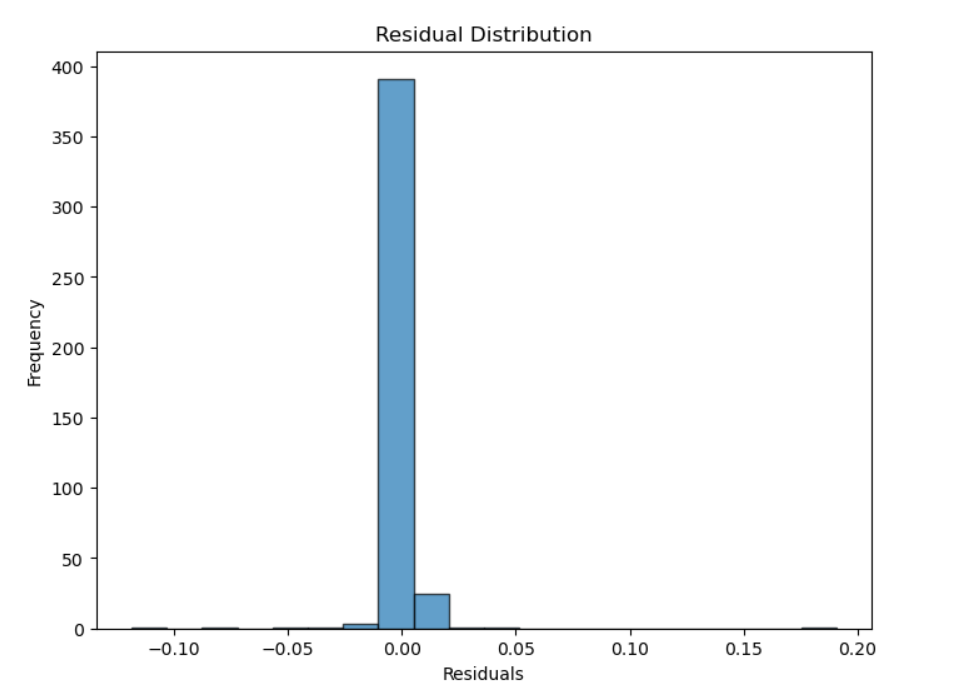
## 7. Visual Analysis

### Actual vs. Predicted Plot



Interpretation of Actual vs. Predicted plot here.

### Residual Distribution



Interpretation of Residual Distribution plot here.