

Model Comparison and Selection Strategy

To predict interest rates (X1), we evaluated Lasso Regression against a Random Forest Regressor. The Lasso model (Baseline) offered the pros of high interpretability and efficient feature selection (eliminating 9 redundant features), but its major con was its inability to model non-linear relationships, resulting in a performance plateau at a Test RMSE of 1.7832.

Conversely, while the Random Forest came with the cons of higher computational cost and "grey box" complexity, it provided the significant pro of capturing high-order interactions and handling noise robustly. By employing a "greedy" strategy with all 39 features and optimizing via RandomizedSearchCV (max_depth=25, n_estimators=300), the Random Forest achieved a superior Test RMSE of 1.3717. This ~23% reduction in error justifies the trade-off in model complexity and computational time for the final holdout prediction.