Group 3: Predicting WAR

Gioia Bonanno-Garcia, Ari Crumley, Vincent West

2025-12-03

Overview

► Goal: build a model to predict WAR (Wins Above Replacement)

► WAR measures a player's total contribution compared to a replacement-level player

► Interpreted as the number of additional wins a player adds to a team

Figure 1: War Ratings Chart

Data used

-we love baseball

Models used

- OLS
  - Simplest model used

- LASSO
  - Feature selection

- Boosting
  - Tree based method

OLS Models

- OLS identifies and measures the relationship between a response variable and predictor variables.
- Finds a best-fitting line through a set of data points
- Pros: Convenient, accurate regression results for linearly related data
- Cons: May be too simplistic for real world examples, assumptions of Linear Regression

OLS Metrics Plot

LASSO Models

- ▶ LASSO models perform regularization (L1), which shrinks some coefficients to exactly zero
  - ▶ Essentially feature selection
- ▶ Pros: Produces a more interpretative model, prevents over fitting
- ▶ Cons: LASSO performs poorly when predictors are highly correlated

LASSO metrics plot

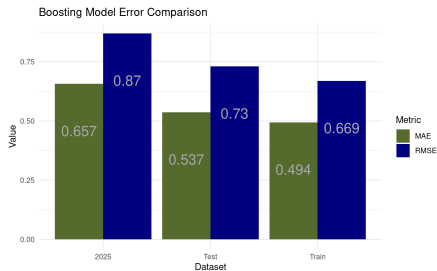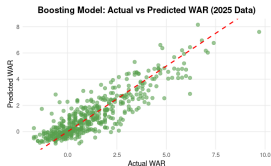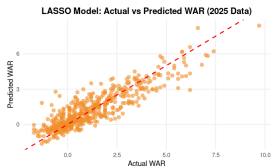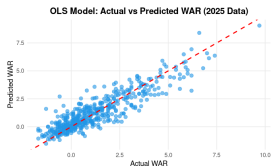- Metrics ran on split training data
- Good performance?

Boosting Models

▶ Boosting grows trees sequentially using information from previously grown trees
  ▶ Each tree fit on a modified version of the original data set

▶ Pros: High predictive accuracy and captures complex, nonlinear relationships automatically.

▶ Cons: Prone to overfitting and requires careful tuning of hyperparameters to perform well.
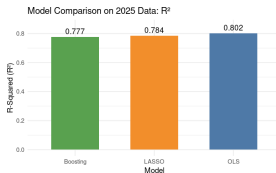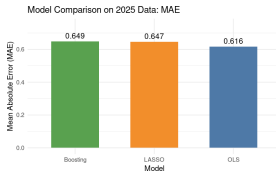
Boosting metrics plot

- Metrics ran on split training data
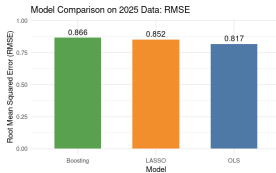

Boosting Model Error Comparison

Model Comparison: OLS, LASSO, and Boosting

OLS Model: Actual vs Predicted WAR (2025 Data)

LASSO Model: Actual vs Predicted WAR (2025 Data)

Boosting Model: Actual vs Predicted WAR (2025 Data)

Model Comparison: RMSE, MAE, and R²

Model Comparison on 2025 Data: RMSE

Model Comparison on 2025 Data: MAE

Model Comparison on 2025 Data: R²

Player	Prediction	Actual
Aaron Judge	9.051733	9.7
Hunter Goodman	3.174746	3.7
Michael Toglia	-1.032315	-1.7
Bobby Witt Jr.	6.125421	7.1