

Group 3: Predicting WAR

Gioia Bonanno-Garcia, Ari Crumley, Vincent West

2025-12-03

- ▶ 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

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

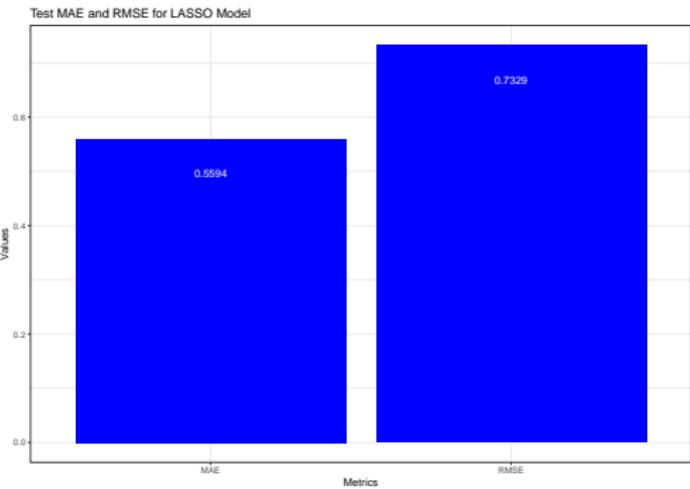


LASSO Models

- ▶ 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 metrics plot

- ▶ Metrics ran on split training data
- ▶ Good performance?

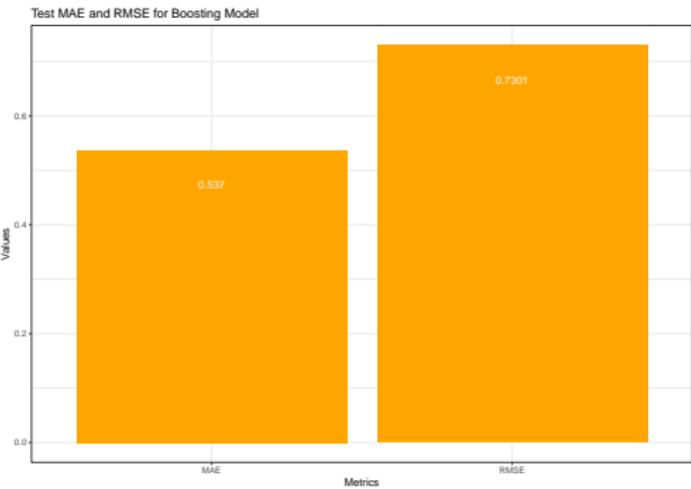


Boosting Models

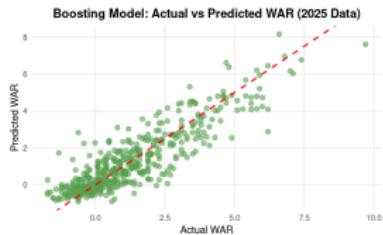
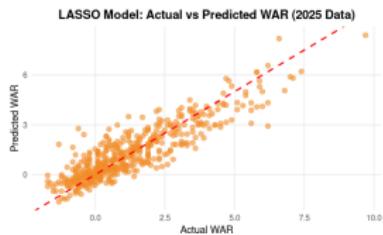
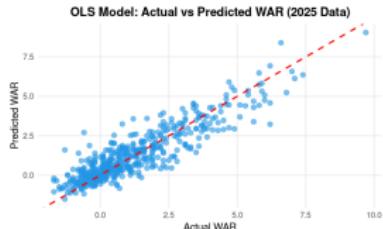
- ▶ Boosting grows trees sequentially using information from previously grown trees
 - ▶ Each tree fit on a modified version of the original data set

Boosting metrics plot

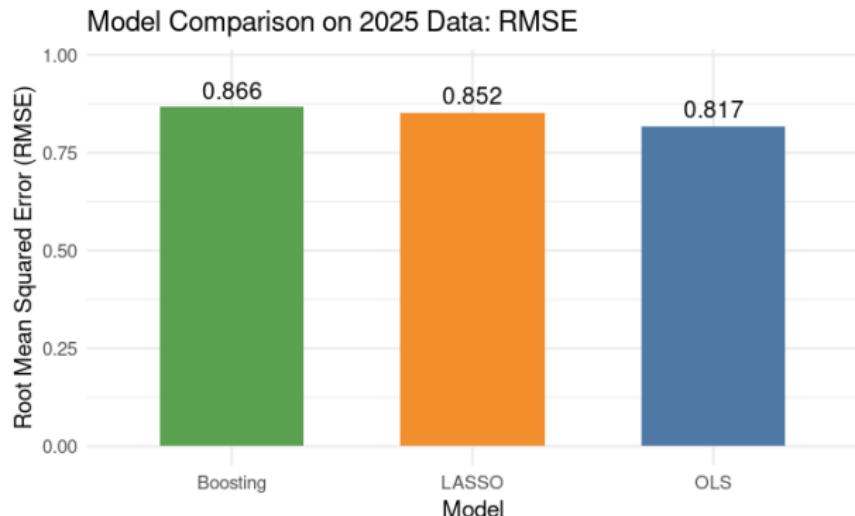
- ▶ Metrics ran on split training data



Model Comparison: OLS, LASSO, and Boosting



Model Comparison: RMSE, MAE, and R²



Model Comparison on 2025 Data: MAE

