Draft & Combine

By: Pender Fitz Gerald, Matilda Soric, Edward Ortiz

Data Exploration

What is the NFL Draft?

- Teams take turns selecting collegiate players
- The first draft was held in 1936
- Currently, the draft consists of seven rounds
- NFL Combine
    - Opportunity for teams to scout a prospect’s raw talent

Goals

- Create a model that predicts the round and/or pick number for each player in the NFL Draft
- Find/determine the correlation between the NFL Combine results and where the player is selected

Our Data

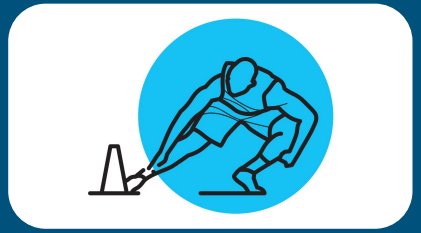
- NFL Draft & Combine data obtained from Kaggle
  - Info from 1993 to 2019 (Over 10,000 players)
  - Combine measurements and personal information
    - Ex. Full Name, Home State, College, Weight, Height, 40-Yard Dash, 20-Yard Shuttle, Vertical Jump…

Cleaning the Data

- Dropped unnecessary columns from both datasets
    - Ex. ‘teamId', 'dob', 'playerProfileUrl', 'homeCountry', 'highSchool', 'hsCity'...
- Merged Combine & Draft datasets on ‘PlayerId’ column
- Filled Null with: ‘ ‘ or 0, depending on data type
- Label ecoded ‘draftTeam’, ‘college’, ‘homeState’ columns

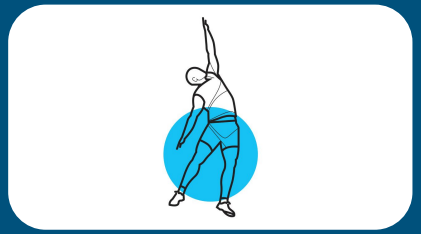
Combine Events

40-Yard Dash

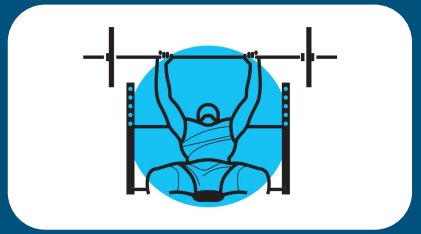


20-Yard Shuttle

Broad Jump



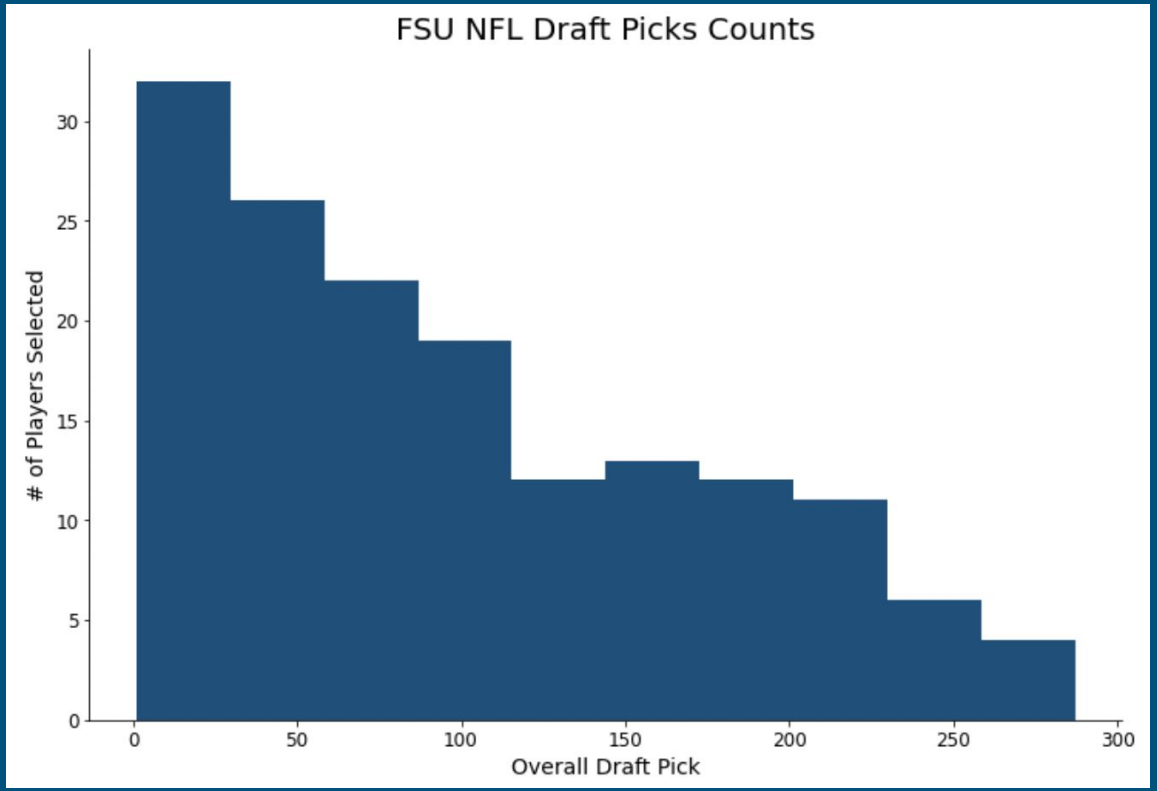
Vertical Jump



Bench Press

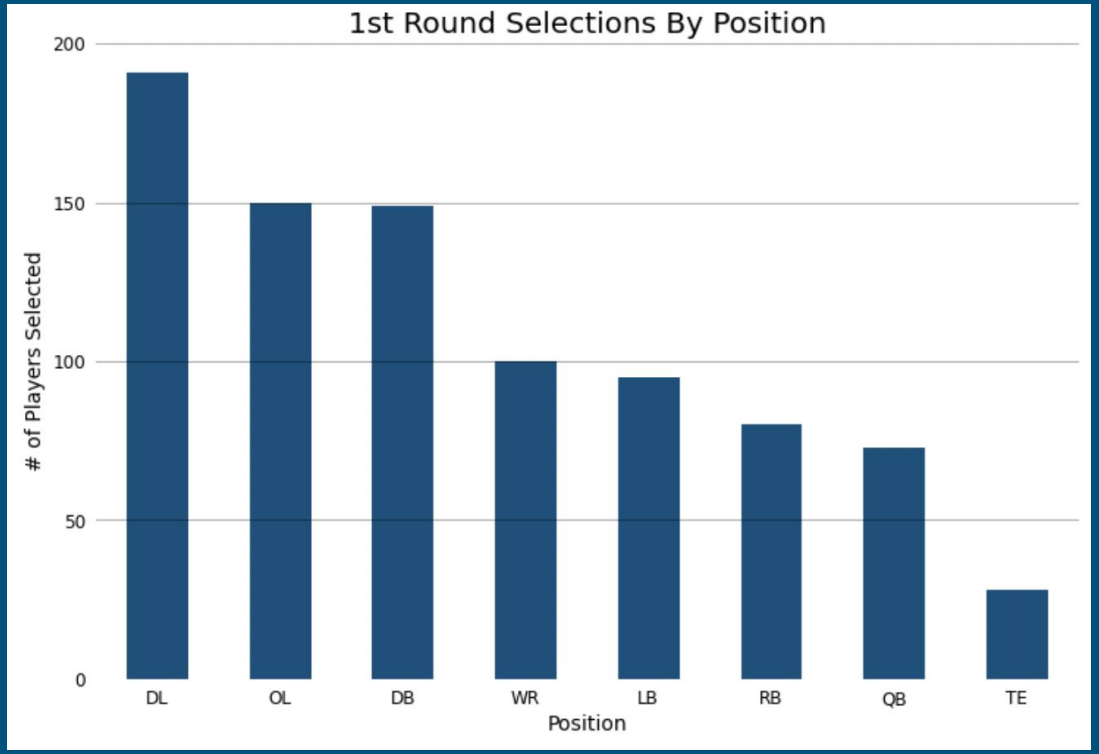
Most Drafted College

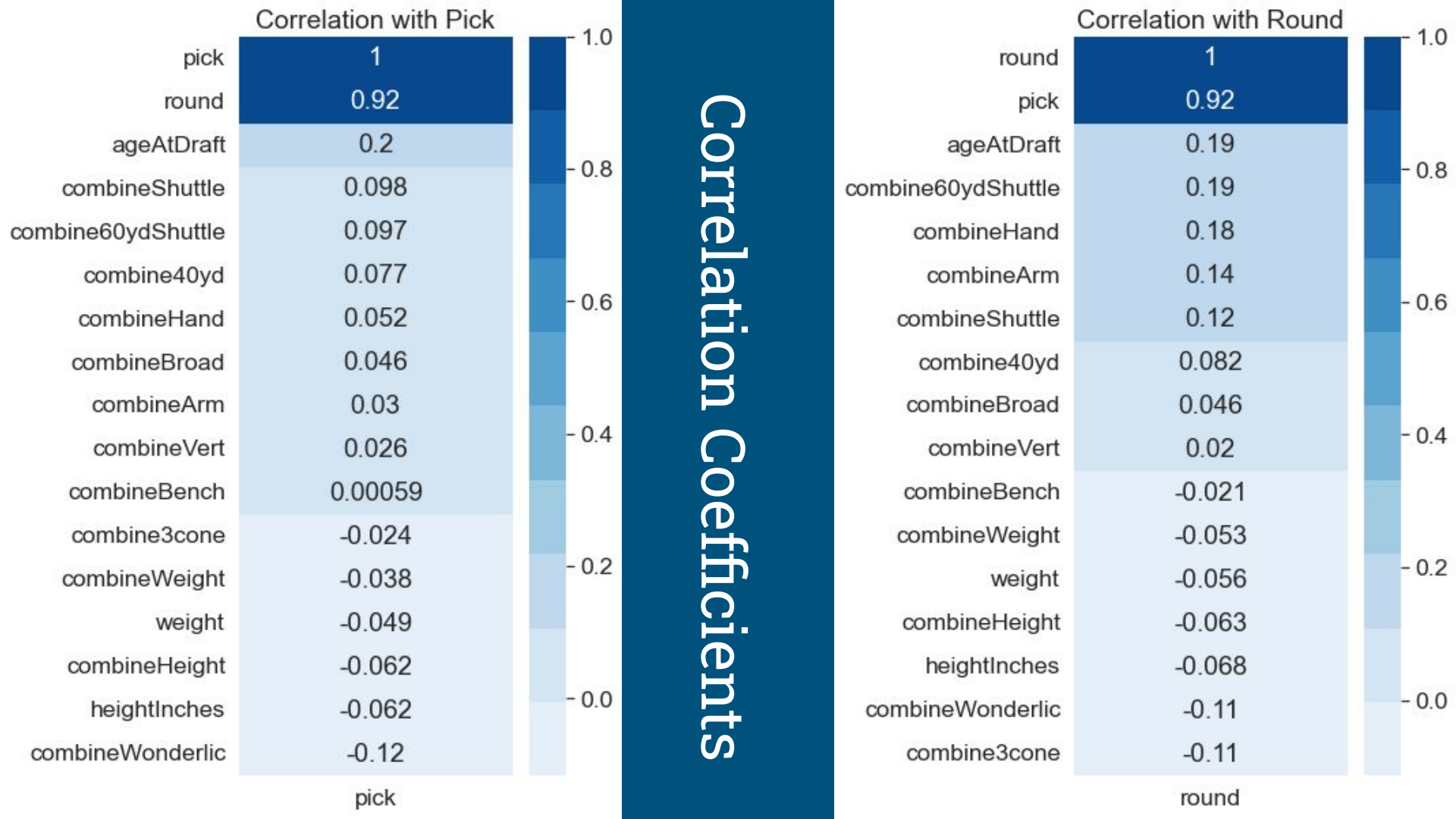
- Florida State
  - 160 drafted players
  - 33 R1 selections



Top Positions Selected in First Round

- Defensive Line (DL) has the overall most picked within first round
- QB is 7th-most selected in first round, but most common 1st overall pick



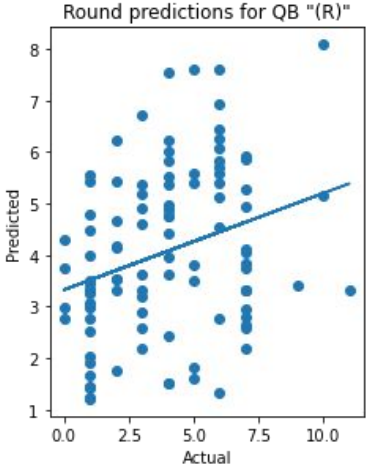

Correlation with Pick
pick 1
round 0.92
ageAtDraft 0.2
combineShuttle 0.098
combine60ydShuttle 0.097
combine40yd 0.077
combineHand 0.052
combineBroad 0.046
combineArm 0.03
combineVert 0.026
combineBench 0.00059
combine3cone -0.024
combineWeight -0.038
weight -0.049
combineHeight -0.062
heightInches -0.062
combineWonderlic -0.12
pick
1.0
0.8
0.6
0.4
0.2
0.0
Correlation Coefficients
Correlation with Round
round 1
pick 0.92
ageAtDraft 0.19
combine60ydShuttle 0.19
combineHand 0.18
combineArm 0.14
combineShuttle 0.12
combine40yd 0.082
combineBroad 0.046
combineVert 0.02
combineBench -0.021
combineWeight -0.053
weight -0.056
combineHeight -0.063
heightInches -0.068
combineWonderlic -0.11
combine3cone -0.11
round
1.0
0.8
0.6
0.4
0.2
0.0

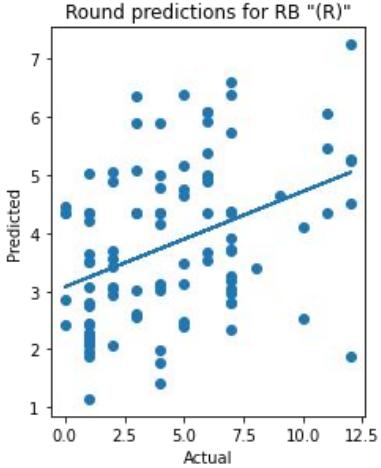
Models

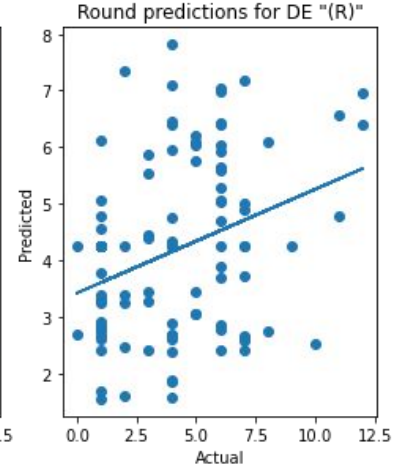
Model for Round Predictictions

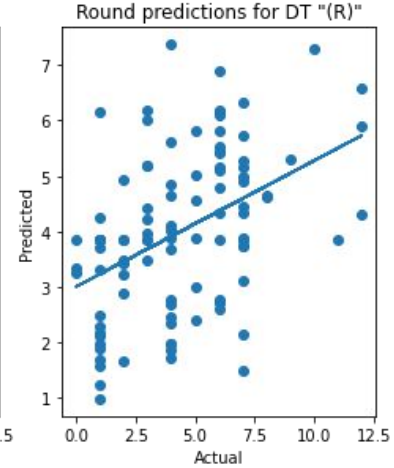
- Columns used :'combine40yd', 'combineVert', 'combineBench', 'combineShuttle', 'combineBroad', 'combine3cone'
- Nearest Neighbors, Random Forest, Decision Tree and XGBoost - Regressors -> average prediction
- Nearest Neighbors, Random Forest, Decision Tree and XGBoost - Classifier -> predictions mode
- Hyperparameter tuning using GridSearchCV

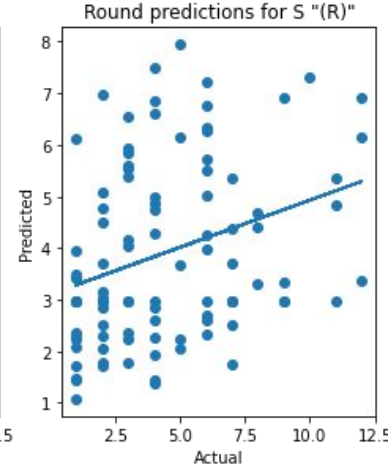
Models Comparison

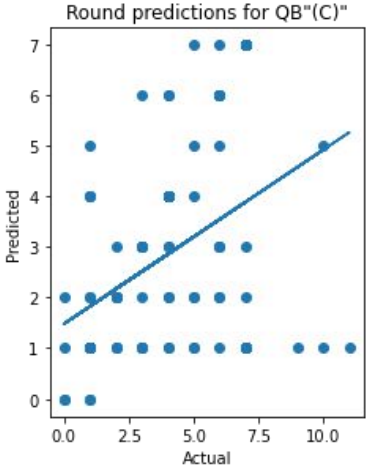

Round predictions for QB "(R)"
Predicted
Actual

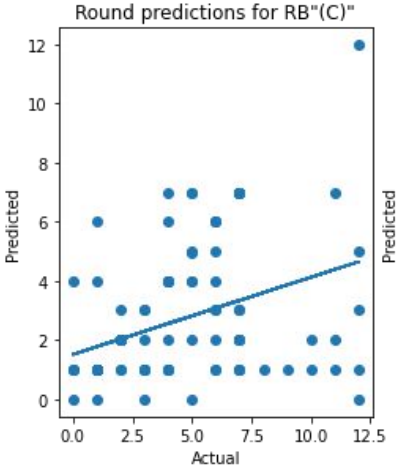

Round predictions for RB "(R)"
Predicted
Actual

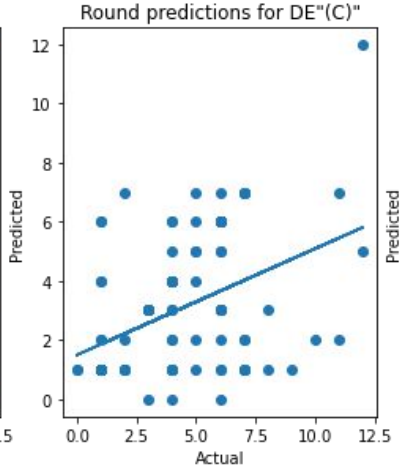

Round predictions for DE "(R)"
Predicted
Actual

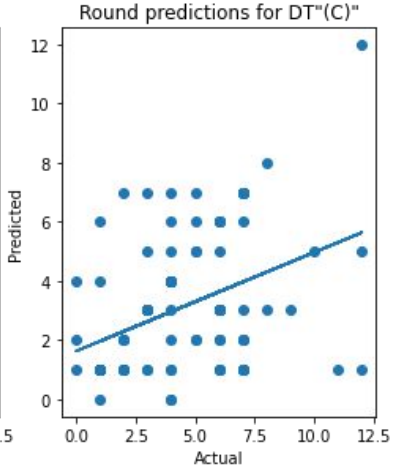

Round predictions for DT "(R)"
Predicted
Actual

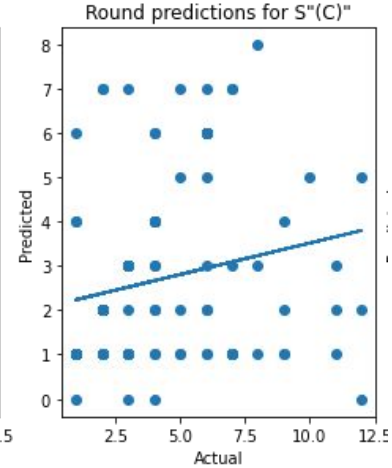

Round predictions for S "(R)"
Predicted
Actual


Round predictions for QB"(C)"
Predicted
Actual


Round predictions for RB"(C)"
Predicted
Actual

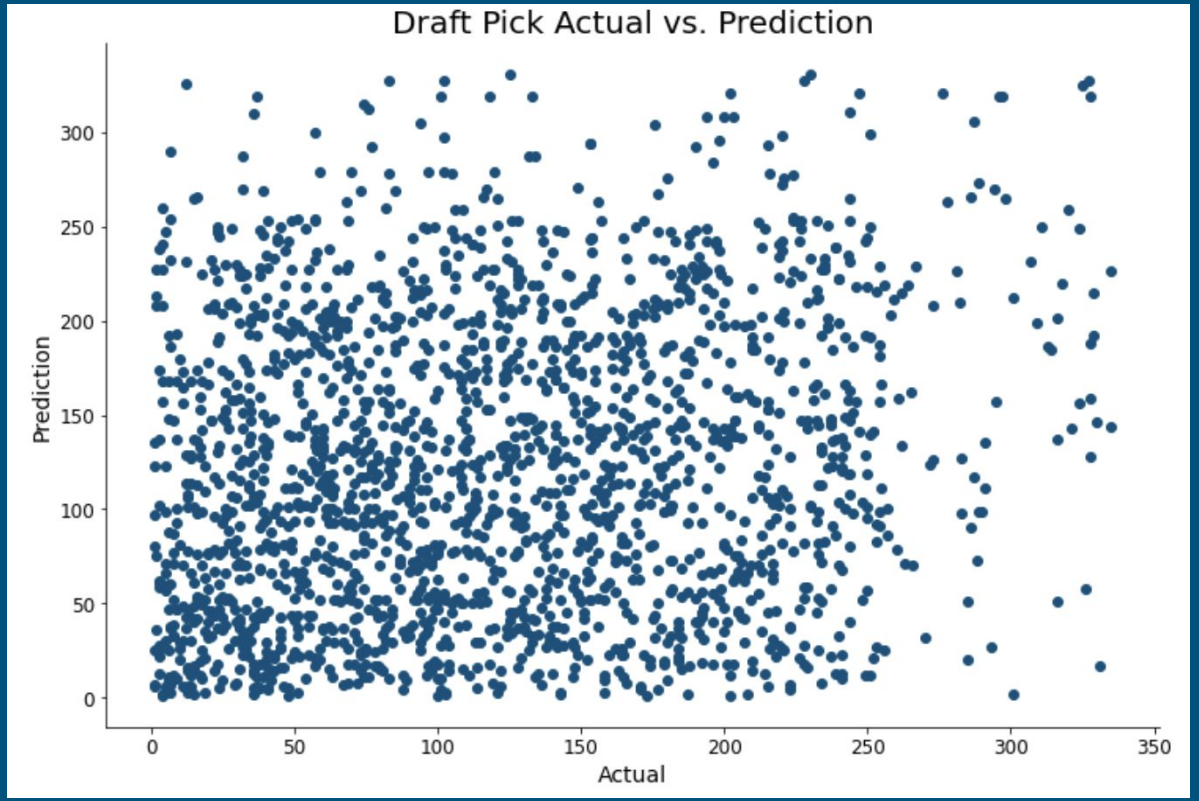

Round predictions for DE"(C)"
Predicted
Actual


Round predictions for DT"(C)"
Predicted
Actual


Round predictions for S"(C)"
Predicted
Actual

Model for Overall Pick

- Events
  - 40-Yard Dash
  - Vertical Jump
  - Bench Press
  - 20-Yard Shuttle
  - Broad Jump
- Measurements
  - Height
  - Weight
  - Hand Size
  - Arm Length
- DT Regressor



Model Challenges

Combine Records

- 40-Yard Dash
  - Trindon Holliday - 2010 (R6): 4.21s
- 20-Yard Shuttle
  - Kevin Kasper - 2001 (R6): 3.73s
- Broad Jump
  - Byron Jones - 2015 (R1): 147 in.
- Vertical Jump
  - Gerald Sensabaugh - 2005 (R5): 46 in.
- Bench Press
  - Stephen Paea - 2011 (R2): 49 of 225 lbs.

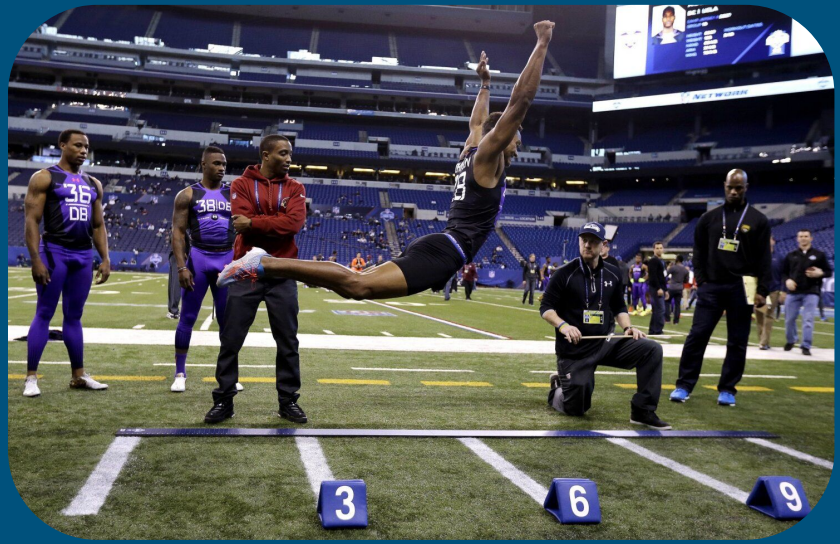


Photo by Julio Cortez - AP

Technical Skill vs. Raw Talent

- Raw Talent
  - Speed (40-Yard Dash)
  - Strength (Bench Press)
  - Explosiveness (Broad & Vertical)
  - Agility (20-Yard Shuttle)
- Technical Ability
  - Football specific
  - Proneness to injury
  - IQ
  - Many more...

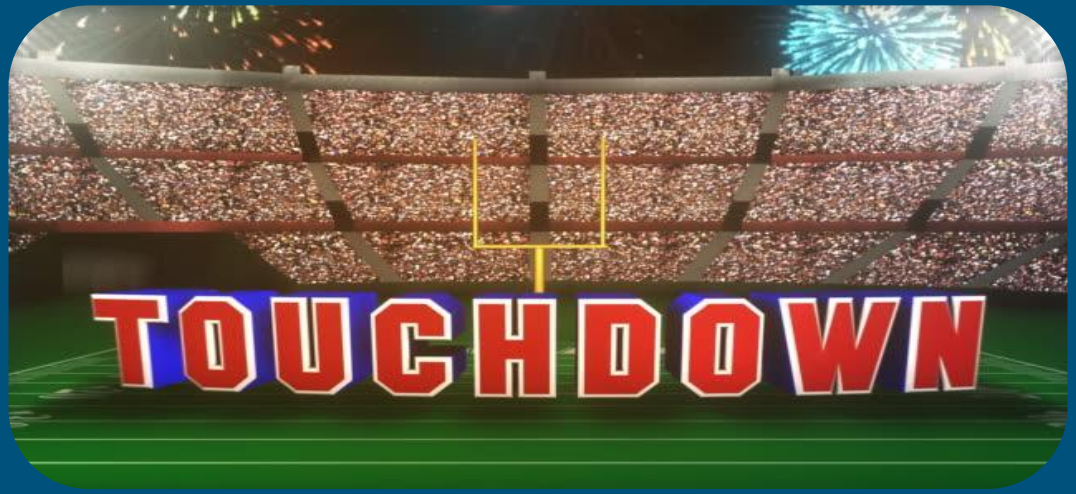
Photo by Brian Spurlock - USA Today

Other Model Challenges

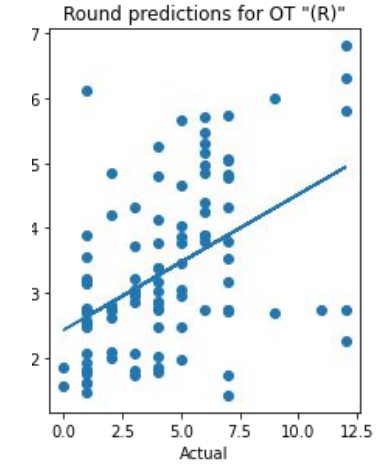
- Not predicting rounds above 8 (in most cases)
  - Last draft with more than 8 rounds was in 1992
- Unbalanced data
  - Small counts of rounds above 8 (5 % of our data)
  - 44% of picks were <= 100
  - 1% of picks were > 300
- Low correlation between variables

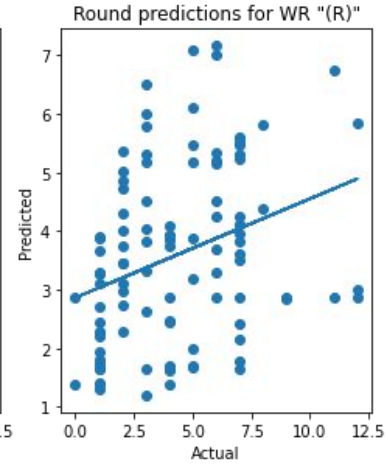
Conclusion

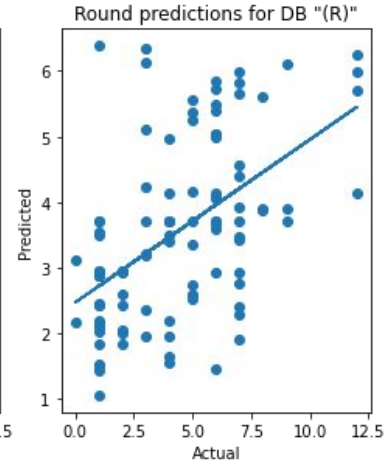
- Combine results are best used in conjunction with other information about a prospect
- A model that has access to measurements of technical skill would most likely perform much better

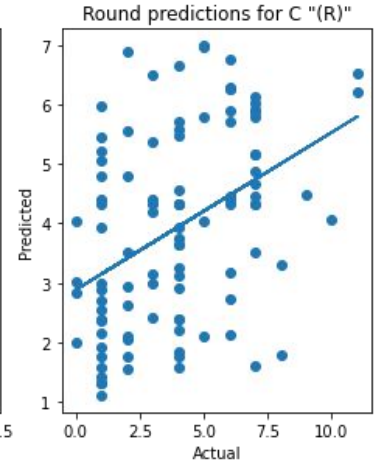


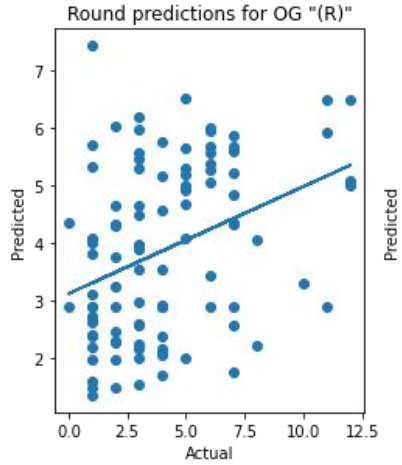
Questions?

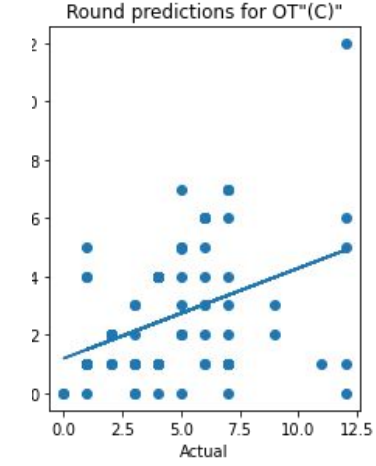

Round predictions for OT "(R)"
Actual

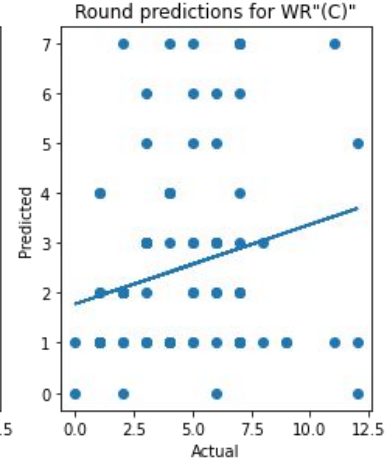

Round predictions for WR "(R)"
Predicted
Actual

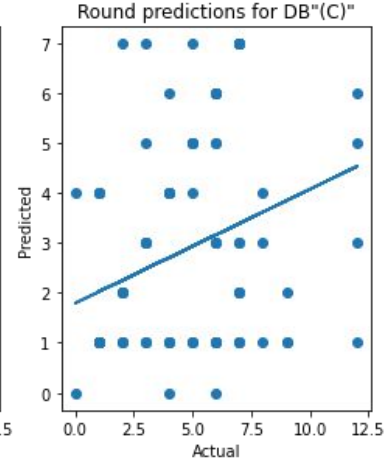

Round predictions for DB "(R)"
Predicted
Actual

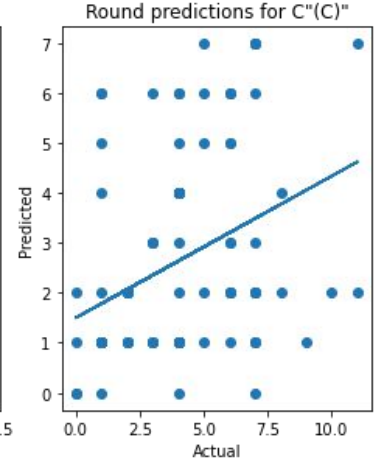

Round predictions for C "(R)"
Predicted
Actual

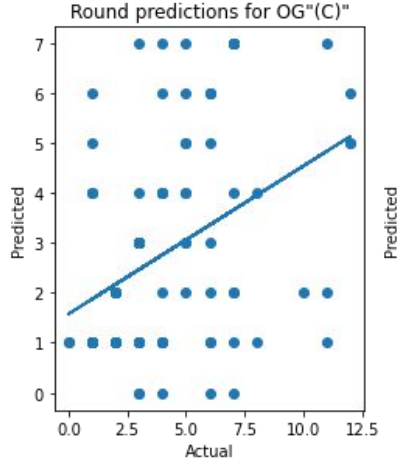

Round predictions for OG "(R)"
Predicted
Predicted
Actual

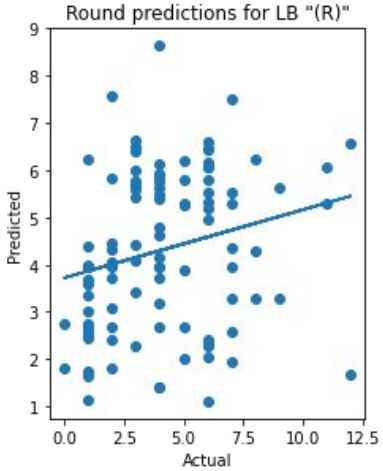

Round predictions for OT"(C)"
Actual

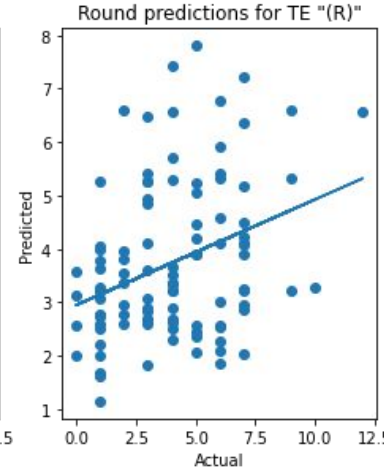

Round predictions for WR"(C)"
Predicted
Actual

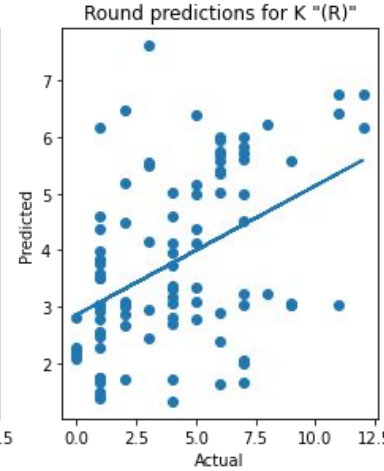

Round predictions for DB"(C)"
Predicted
Actual

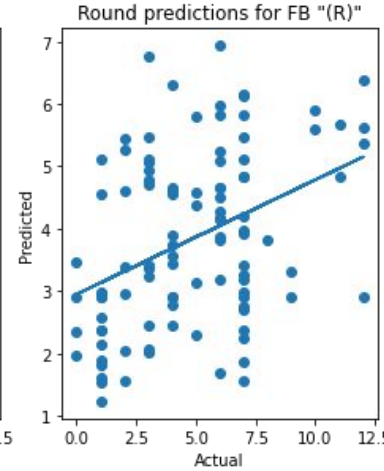

Round predictions for C"(C)"
Predicted
Actual

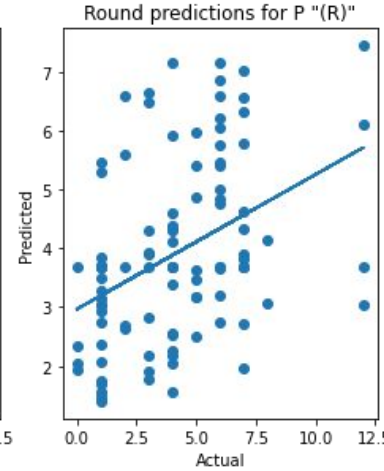

Round predictions for OG"(C)"
Predicted
Predicted
Actual

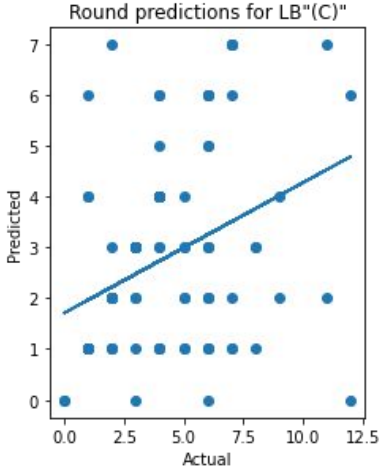

Round predictions for LB "(R)"
Predicted
Actual

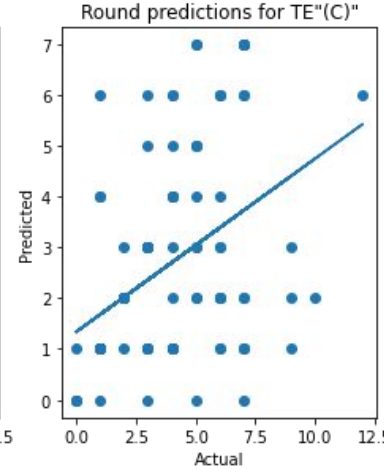

Round predictions for TE "(R)"
Predicted
Actual

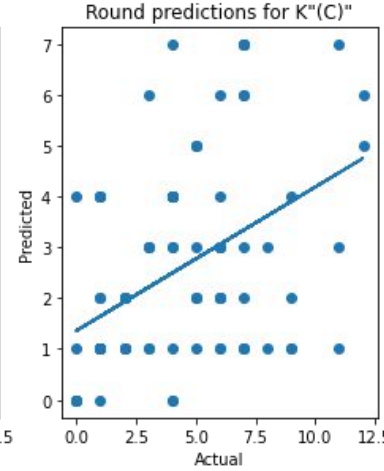

Round predictions for K "(R)"
Predicted
Actual

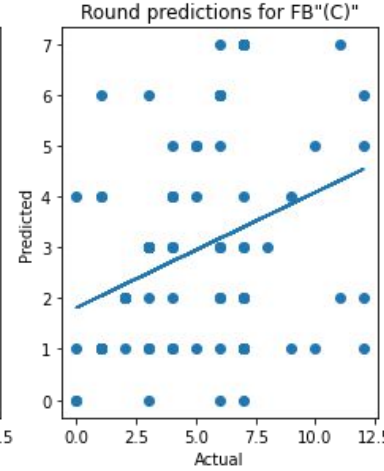

Round predictions for FB "(R)"
Predicted
Actual

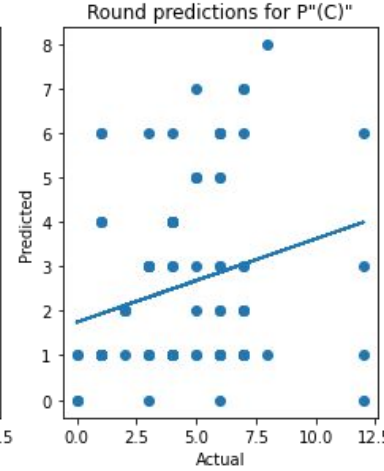

Round predictions for P "(R)"
Predicted
Actual

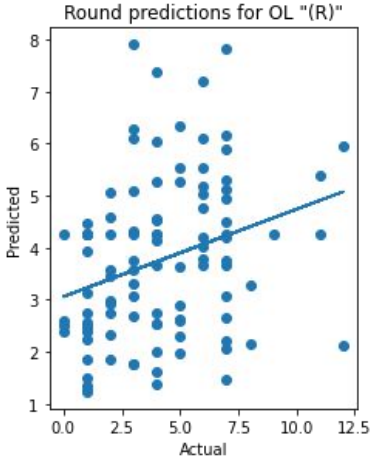

Round predictions for LB"(C)"
Predicted
Actual

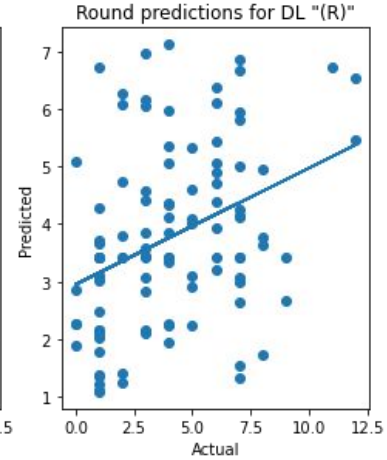

Round predictions for TE"(C)"
Predicted
Actual

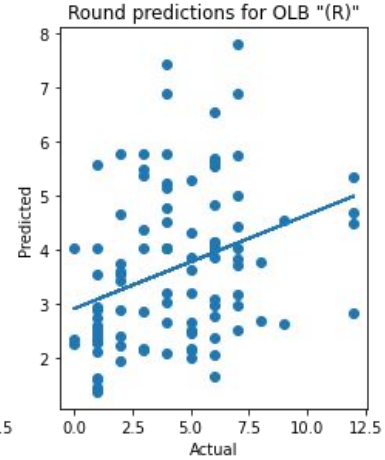

Round predictions for K"(C)"
Predicted
Actual

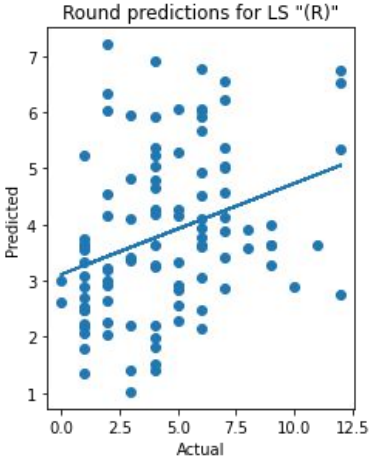

Round predictions for FB"(C)"
Predicted
Actual

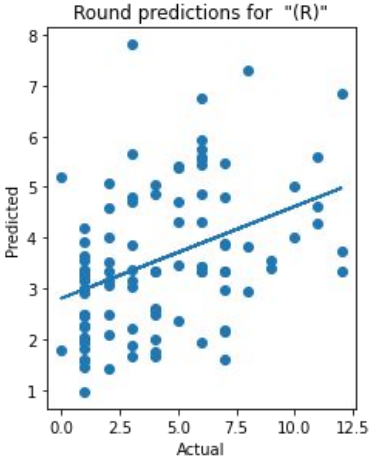

Round predictions for P"(C)"
Predicted
Actual

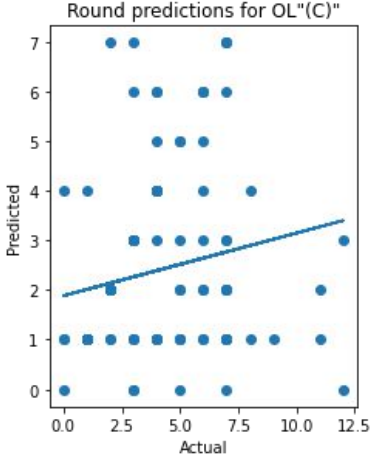

Round predictions for OL "(R)"
Predicted
Actual

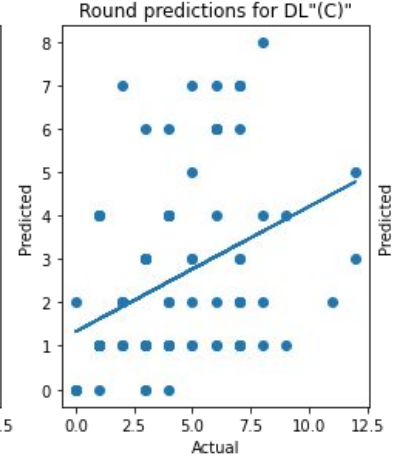

Round predictions for DL "(R)"
Predicted
Actual

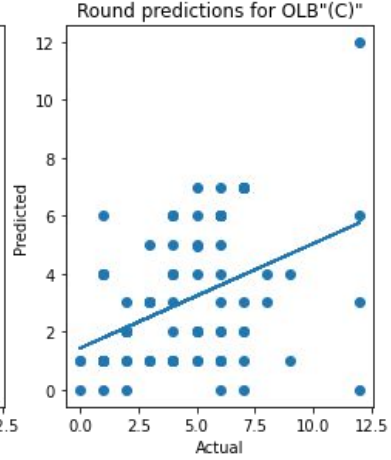

Round predictions for OLB "(R)"
Predicted
Actual

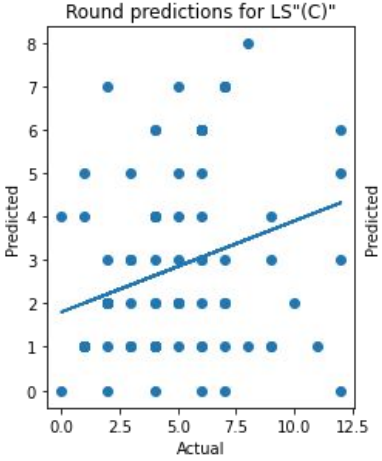

Round predictions for LS "(R)"
Predicted
Actual

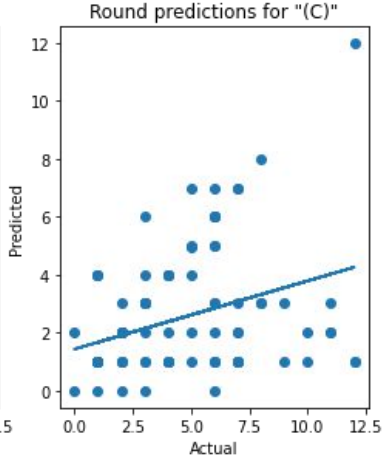

Round predictions for "(R)"
Predicted
Actual


Round predictions for OL"(C)"
Predicted
Actual


Round predictions for DL"(C)"
Predicted
Actual


Round predictions for OLB"(C)"
Predicted
Actual


Round predictions for LS"(C)"
Predicted
Actual


Round predictions for "(C)"
Predicted
Actual