2020 NFL WR Draft Class Model

By: Ruslan Davtian

Outline

- Research Question
- Data Collection
- Methodology
  - NFL WR Performance Metric(s)
  - Factors of Interest
  - Modeling Approach
  - Predictions
- Results
- Recommendations/Improvements
- Appendix

Research Question

- Can we use data to rank 2020 WR draft candidates?
    1. Use past NFL pre-draft data to build a model to predict NFL WR performance
    2. Use model to make predictions for the 2020 WR class
    3. Interpret Results & understand what factors influence NFL performance

Data Collection

- Scrape Publicly available Pro Football Reference data (2000-2020)
  - Game-by-game and season-by-season statistics
  - NOT play-by-play or player coordinate level data

- College Receiving Stats by Season
  - Games, Receptions, Receiving Yards, TDs, etc.
- NFL Draft Combine
  - Age, College, Height, Weight, Vertical, 40 yd, Broad Jump, 3 Cone, Shuttle
- Pro Receiving Stats by Season
  - Games, Targets, Receptions, Catch %, Receiving Yards, TDs etc.
- Player Drafted List

Methodology

- What years to gather (Last 20 seasons)
- What metric(s) should be used as the dependent/response variable
  - Use one metric vs multiple metrics to rank players
  - Averaged first 3 seasons
- What metrics should be used as independent/predictor variables
  - Anything useful information available before NFL draft
- What model to use?
  - Regression vs. Machine Learning
  - One model vs Many

NFL WR Performance Metric(s)

- Targets Per Game (Num_Tgt_Per_Game)
- Average Catch Percentage (Avg_Catch_Perc)
- Receptions Per Game (Num_Rec_Per_Game)
- Receiving Yards Per Game (Rec_Yds_Per_Game)
- Receiving Yards Per Reception (Rec_Yds_Per_Rec)
- Receiving TDs Per Game (Rec_TD_Per_Game)
- 1st Down Receptions Per Game (First_Down_Rec_Per_Game)
- Receiving Yards Per Target (Rec_Yds_Per_Tgt)

Factors of Interest

- NFL Combine
  - Age, Height, Weight, BMI, 40 YD, Vertical, Broad Jump, 3 Cone, Shuttle

- College Statistics
  - Conference, Power Conference, Number of Seasons,
  - Number of Receptions Per Game, Receiving Yards Per Game, Receiving Yards Per Reception, Receiving TDs Per Game, Number of Plays Per Game
  - Total TDs, Total Receiving Yards
  - Scrimmage Yards Per Play, Scrimmage TDs Per Game

Modeling Approach

- Build 8 machine learning models using the same factors to predict each performance metric

- Make predictions for the 2020 WR draft class for each of the 8 metrics

- Rank each player from best to worst in each metric

- Find average or median rank of all metrics across each player

Predictions (Top 15 of 48)

	Player	MedianRank	Tgt Per G	Avg Catch %	Num Rec Per G	Rec Yds Per G	Rec Yds Per Rec	Rec TD Per G	1st Down Rec Per G	Rec Yds Per Tgt
1	Tyler Johnson	2.5	2	9	2	3	10	9	2	2
2	James Proche	3.5	1	15	1	1	24	10	1	6
3	CeeDee Lamb	4.0	4	38	4	4	3	8	15	1
4	Omar Bayless	5.0	3	18	3	2	9	13	3	7
5	Tee Higgins	5.5	6	47	8	7	1	2	5	3
6	Gabriel Davis	6.0	5	28	6	6	6	7	6	8
7	Jerry Jeudy	8.5	8	31	5	8	5	17	10	9
8	Chase Claypool	8.5	10	34	14	5	7	1	4	10
9	Justin Jefferson	10.0	7	32	9	9	16	19	7	11
10	Bryan Edwards	12.0	12	8	11	12	27	16	8	14
11	Quez Watkins	12.5	9	12	13	14	14	15	12	12
12	Michael Pittman	13.0	14	44	12	10	17	5	13	13
13	K.J. Hill	14.0	13	1	7	15	48	40	11	27
14	Jauan Jennings	14.0	17	7	10	11	34	31	9	19
15	Denzel Mims	16.5	16	46	18	17	8	4	23	5

Results (Assuming 2020 Draft Happened)

Predicted (Actual Round Drafted)

1. Tyler Johnson (5th round)
2. James Proche (6th round)
3. **CeeDee Lamb (1st round)**
4. Omar Bayless (Undrafted)
5. **Tee Higgins (2nd round)**
6. Gabriel Davis (4th round)
7. **Jerry Jeudy (1st round)**
8. *Chase Claypool (2nd round)*
9. **Justin Jefferson (1st round)**
10. Bryan Edwards (3rd round)

Actual Draft Order

1. Henry Ruggs III
2. **Jerry Jeudy**
3. **CeeDee Lamb**
4. Jalen Reagor
5. **Justin Jefferson**
6. Brandon Aiyuk
7. **Tee Higgins**
8. Michael Pittman Jr
9. Lavishka Shenault Jr.
10. KJ Hamler

Recommendations/Improvements

- Discrepancies between model and expected mock drafts exist
- Why Tyler Johnson, James Proche, Omar Bayless?
- Bias in chosen metrics to predict? Should it be weighted equally
- Data Improvements
  - More granular level
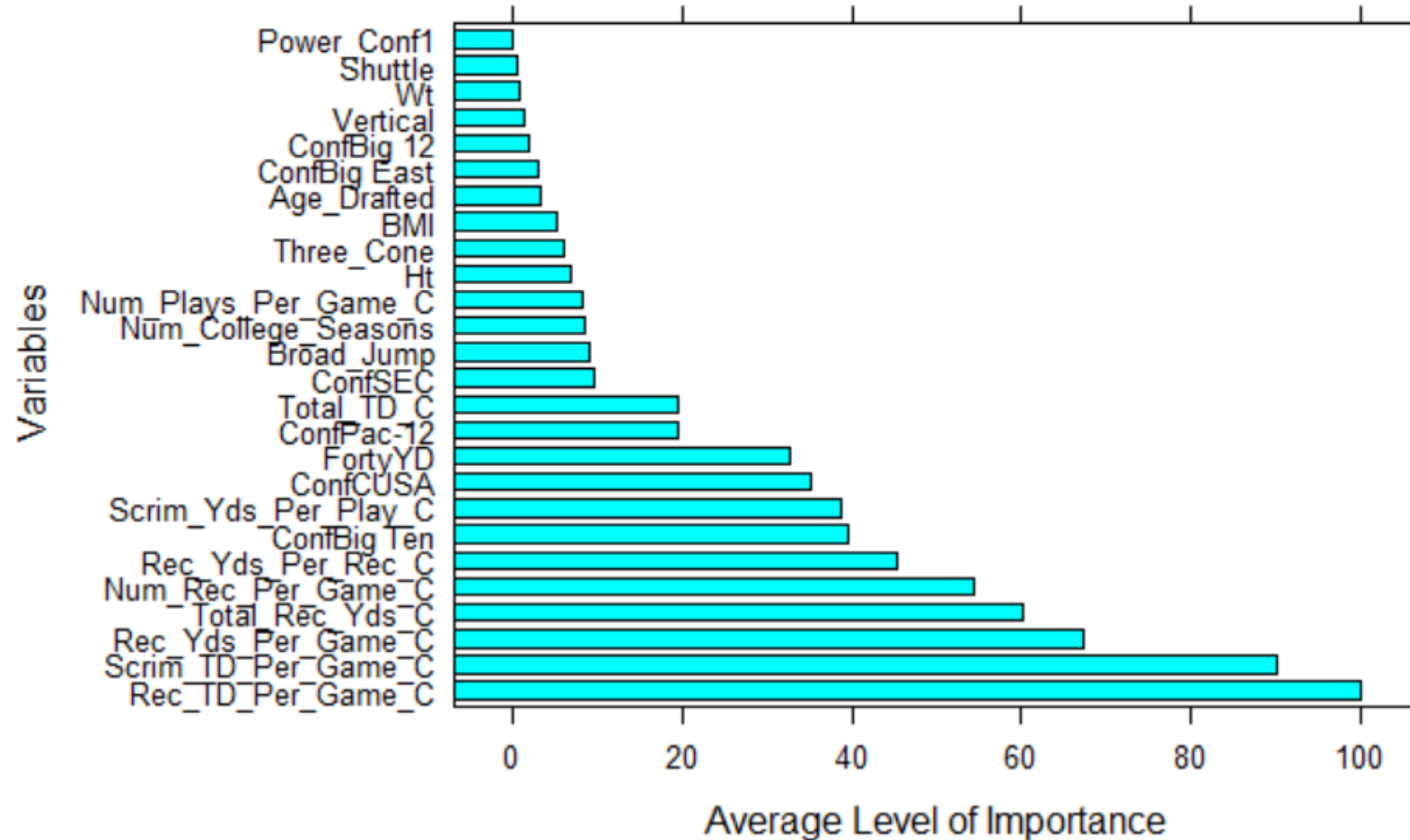- Analytics + Scouts = More informed decision making

Appendix

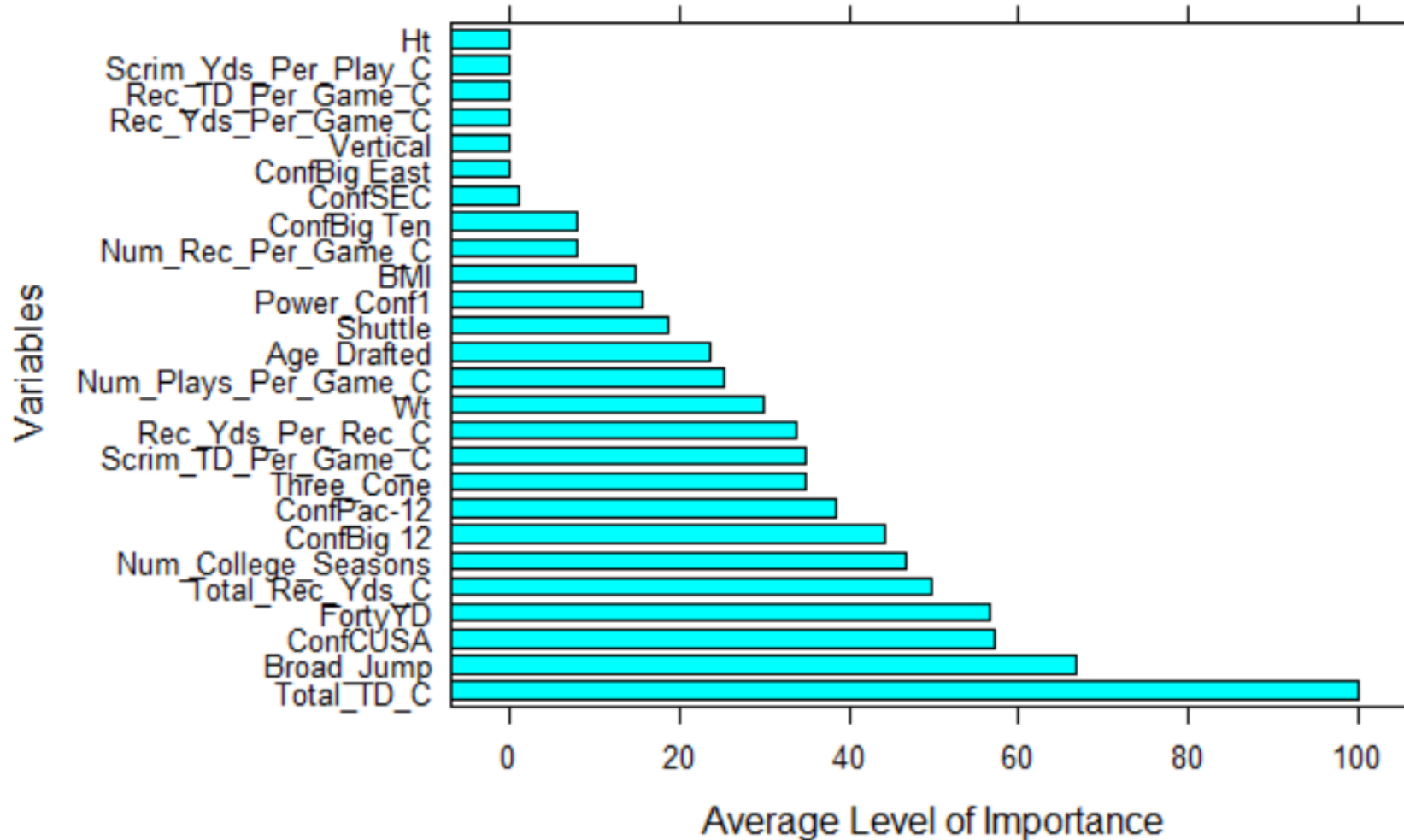**glmnet Variable Importance Predicting Num_Tgt_Per_Game**
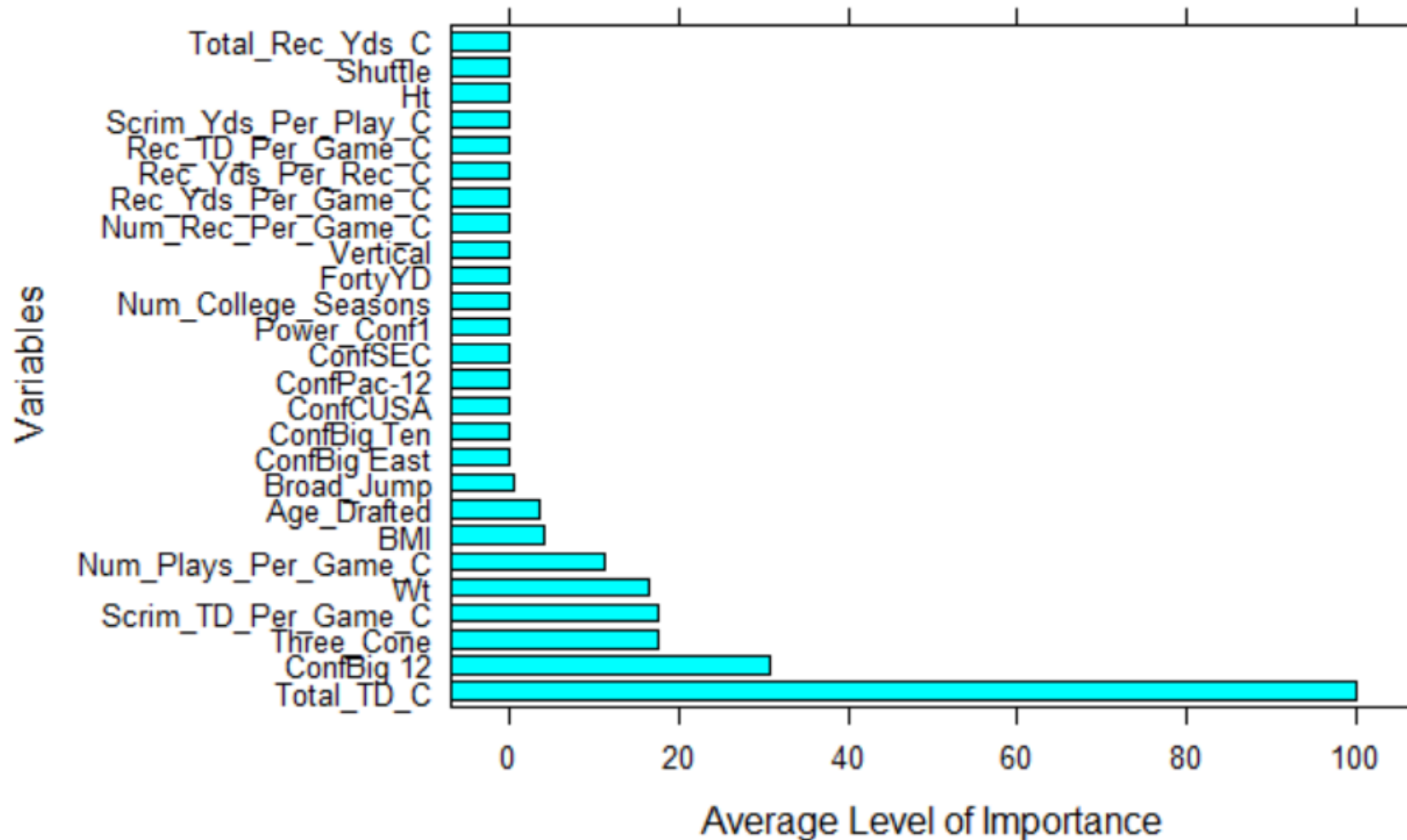
Appendix

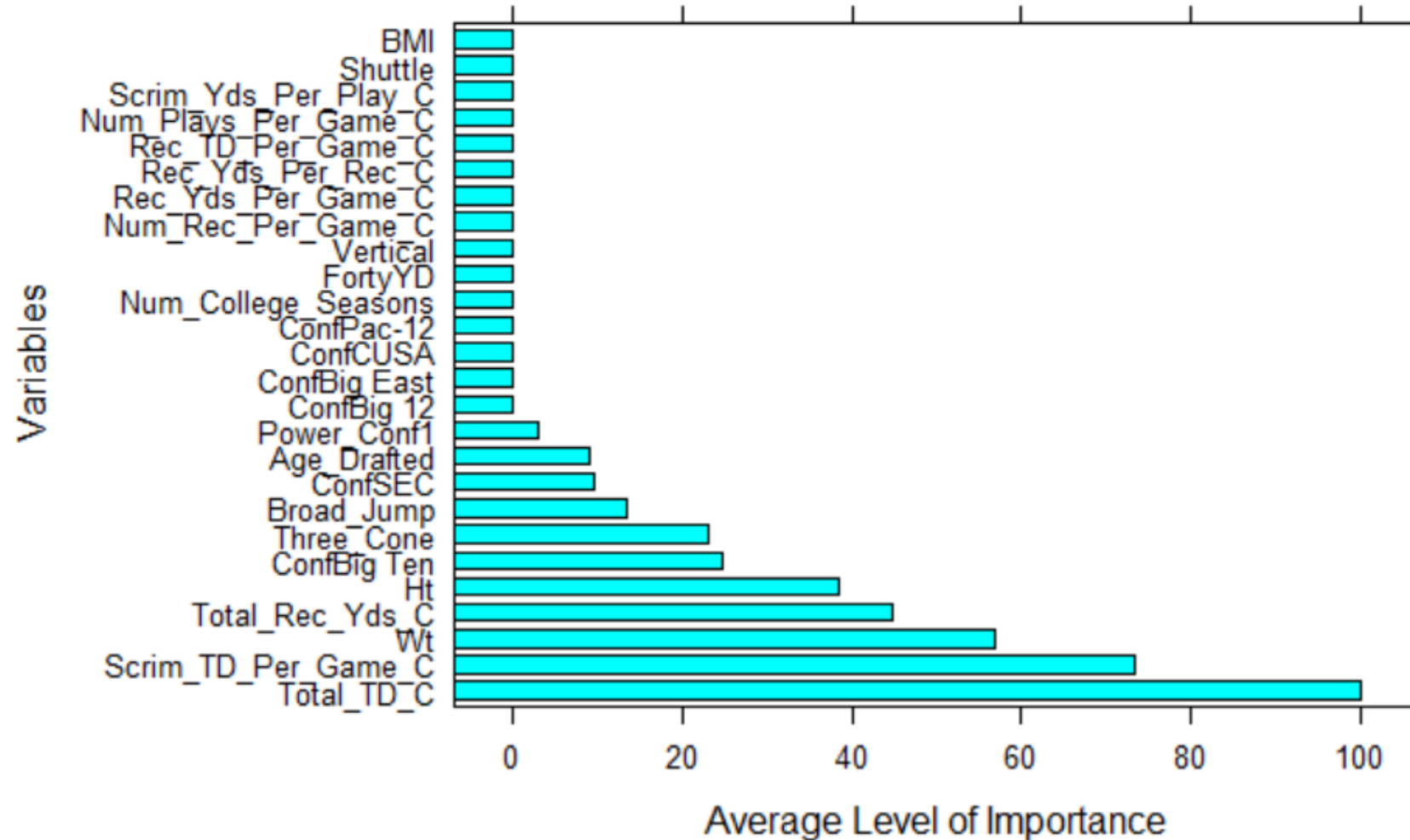**glmnet Variable Importance Predicting Avg_Catch_Perc**

Appendix

glmnet Variable Importance Predicting Num_Rec_Per_Game
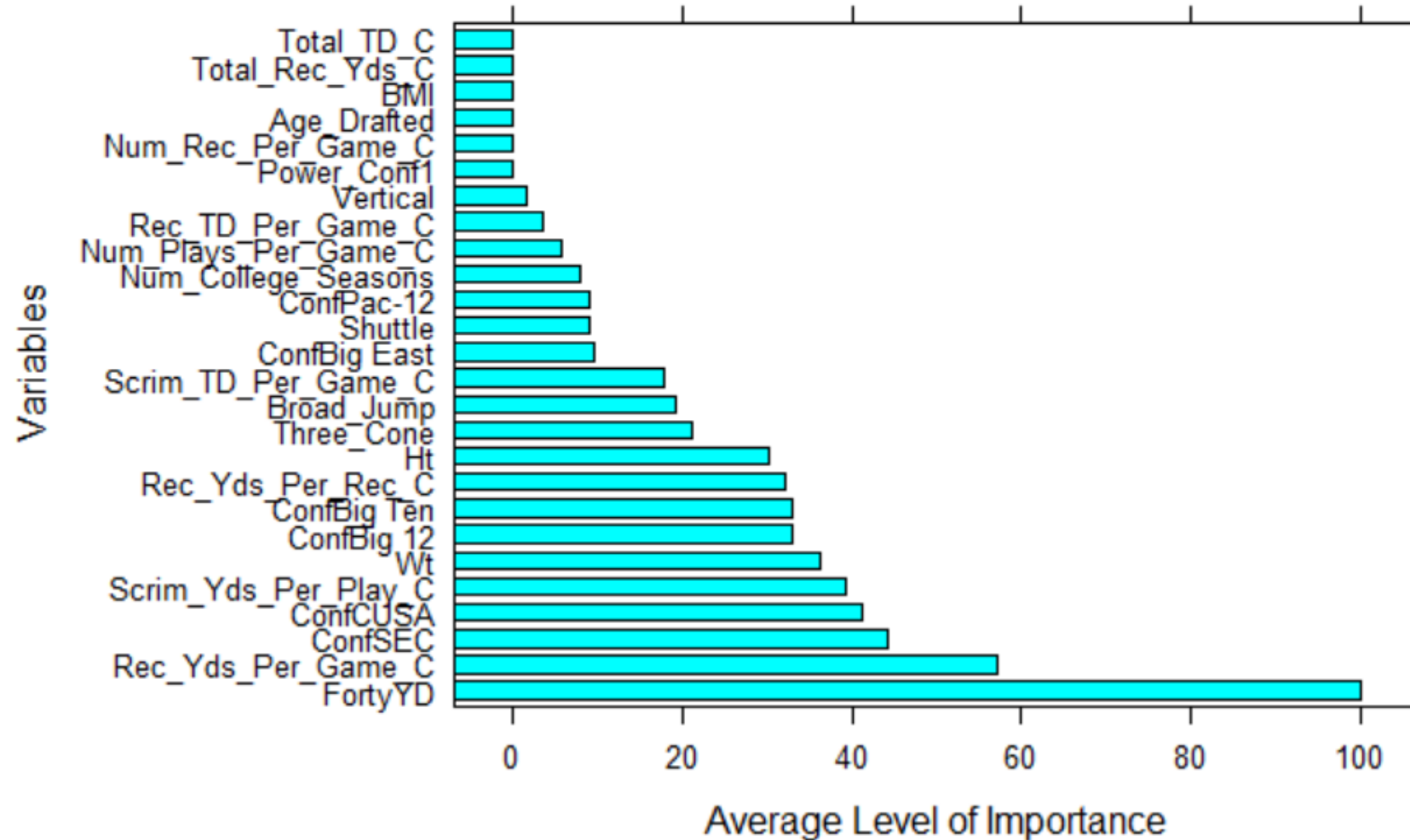
Appendix

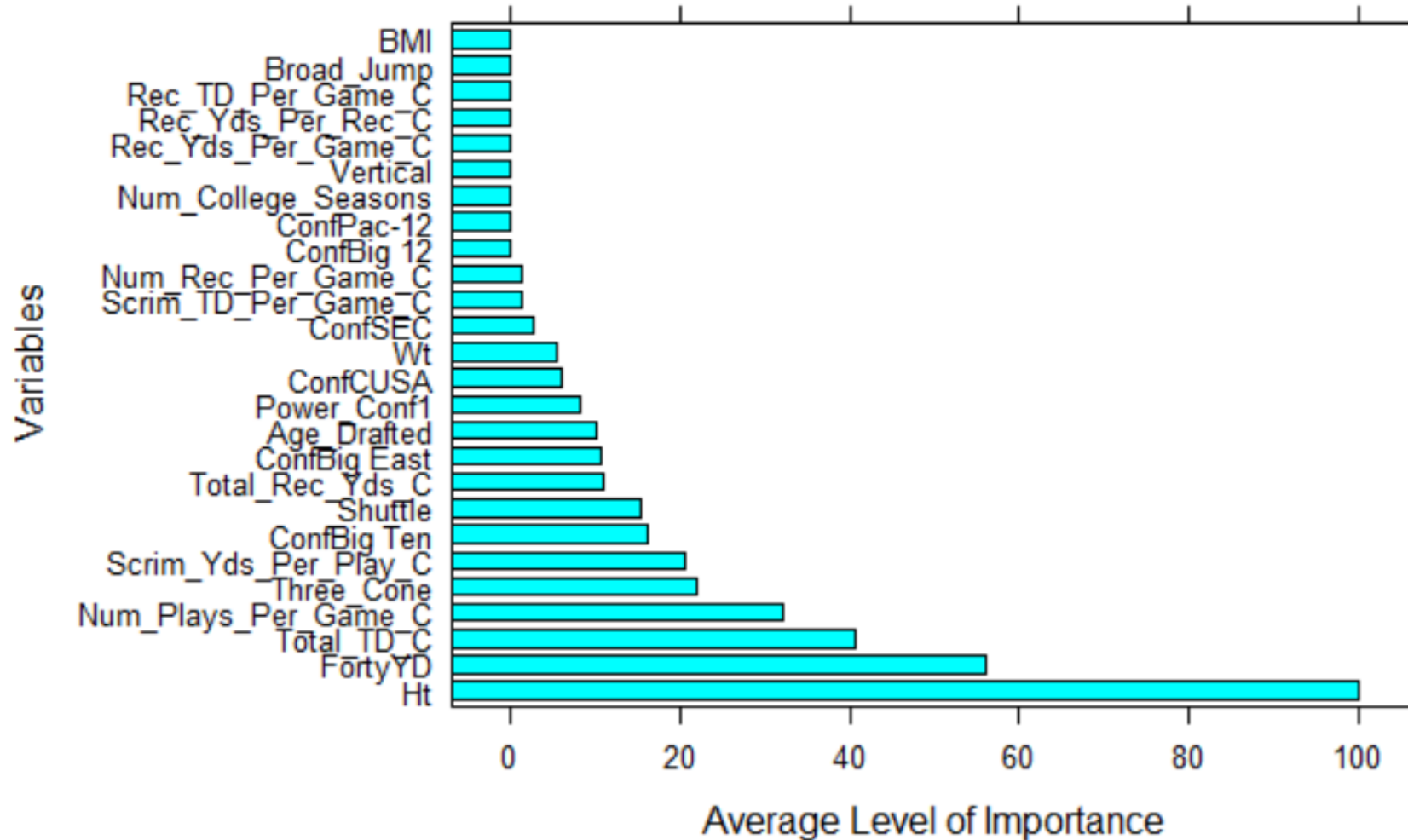glmnet Variable Importance Predicting Rec_Yds_Per_Game

Appendix
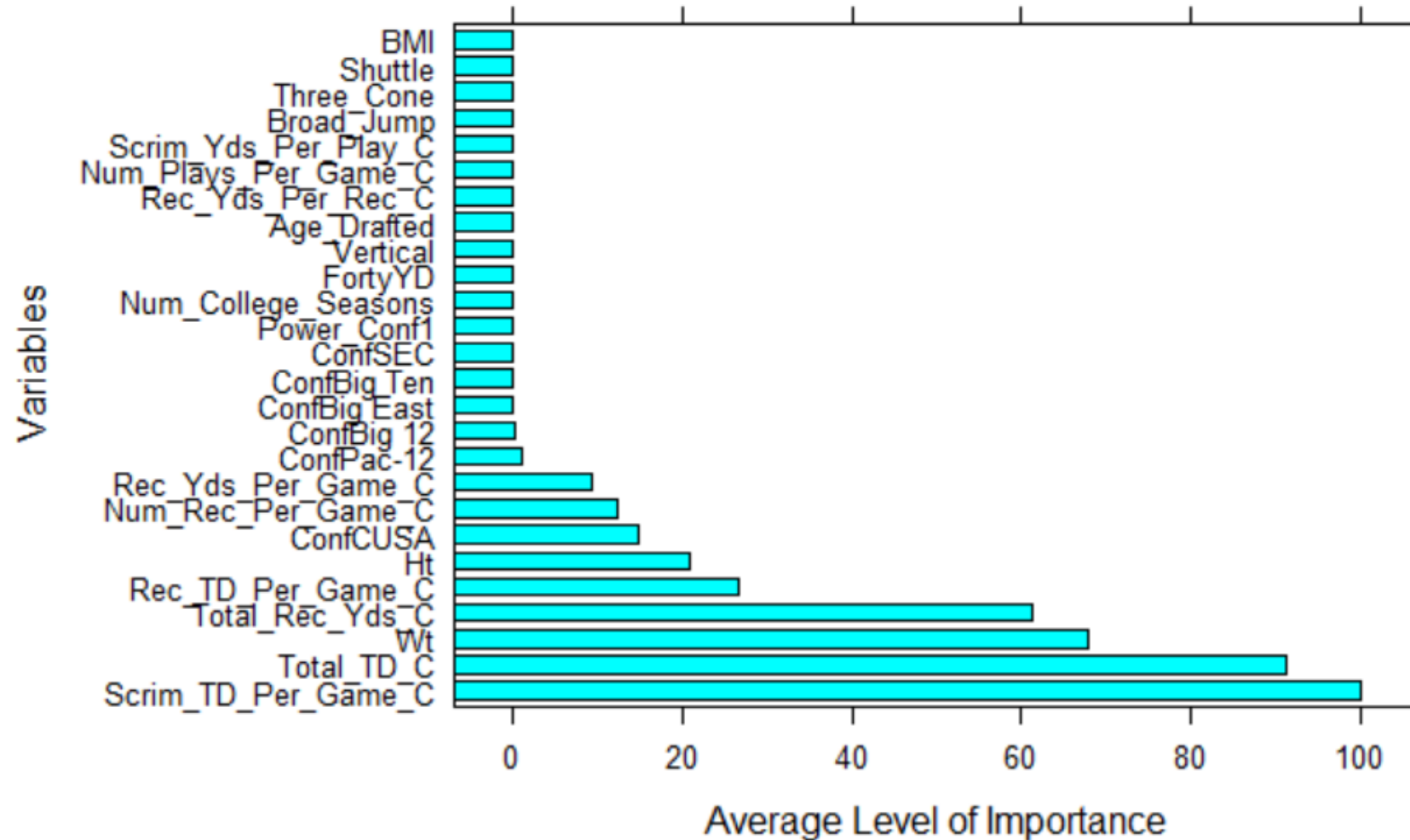

glmnet Variable Importance Predicting Rec_Yds_Per_Rec

Appendix

glmnet Variable Importance Predicting Rec_TD_Per_Game

Appendix

glmnet Variable Importance Predicting First_Down_Rec_Per_Game

Appendix

glmnet Variable Importance Predicting Rec_Yds_Per_Tgt