

Regression Analysis: Pro Football

Predicting performance based on college statistics

By: Eric Patterson

Objective

- Predict professional performance using college stat lines
 - Analyzed the three main skill positions: QB, WR, and RB
- Predictive Models
 - Predict professional performance to the highest extent possible
- Explanatory Models
 - See which stats are the most influential to professional performance

Methodology: Overview

- Used both cumulative college statistics and final year statistics to predict performance
- Predictive models were tested for all regression types
- Explanatory models use Linear Regression (no regularization) to maintain interpretability
- Model uses a 10% confidence threshold for explanatory variables

Methodology: Determining Player Value

- Dependent variable: player's lifetime performance
- QB: passer rating
- Chased lowest MSE
- RB and WR: custom value formula:
$$(Yds + 50 * (TD - 0.5 * Fumbles)) / (0.5 * Games)$$

Wide Receiver Elements Examined

- Receptions
- Yards
- Touchdowns
- Yards per Catch
- Year played (time factor)
- Conference
- Interaction and polynomial terms

Wide Receiver Findings

- Explanatory Model:
 - Career stats at college
 - Yards, Receptions, Touchdowns and playing in a Power Conference
 - 12.2% of Pro performance explained
- Predictive Model:
 - Use only senior year college stats
 - Throw all stats in a Lasso model, plus receptions * conference
 - 12.3% of Pro performance explained

Quick Takeaway

- Playing in a Power Conference is the most important at this skill position

Running Back Elements Examined

- Rushing Attempts
- Yards
- Touchdowns
- Yards per Rush
- Year played
- Conference
- Interaction and polynomial terms

Running Back Findings

- Explanatory Model:
 - Career stats at college
 - Yards, Touchdowns and Attempts * Year
 - 15.7% of Pro performance explained
- Predictive Model:
 - Use only senior year college stats
 - Elastic Net: Yards, Touchdowns and Attempts * Year
 - 16.6% error reduction from explanatory model

Quick Takeaway

- It is becoming more important to get a running back that has more college experience
 - Flies in the face of the position's youth movement

Quarterback Elements Examined

- College passer rating
- Yards
- Touchdowns and Interceptions
- Yards per Throw
- Attempt and Completions
- Completion Percentage
- Year played
- Conference
- Interaction and polynomial terms

Quarterback Findings

- Explanatory Model:
 - Use only senior year college stats
 - Attempts, Completions and Percentage * Year
 - 18% of Pro performance explained
- Predictive Model:
 - Career stats at college
 - Linear Regression with all stats
 - 21.7% error reduction from explanatory model

Quick Takeaway

- It is becoming more important to get a quarterback with accuracy

Overview of Model Explanatory Power

- Variation explained: 1/5th (QB's) to 1/10th (WR's)
 - Importance of player development over raw skill
- Positions with better explanatory models have even better predictive models

Recap of Takeaways

Position	What to Look For:
Quarterbacks	Accuracy/Completion Percentage
Running Backs	More college experience
Wide Receiver	Attended a Power 5 school

Appendix

- Scraped data from football-reference.com