

NBA Draft Prediction

Leveraging college basketball statistics to predict the NBA draft pick Authors: Kajal Tiwary, Austin Eaton, Jackson Piccione, Abby Fremaux

Introduction



Problem to Address

How to predict NBA draft selections using college basketball player statistics

Importance

Teams can develop more comprehensive strategies to acquire the best talent

Background

What: Selection of players to join the NBA When: Annually every summer (June)

Who: Individuals 19 or older

How: Teams with worse records select first

Where: Changes year to year

Methods



Exploratory Analysis

Understand the Landscape

- Top Schools
- Draft pick correlation with key statistics
- Key statistic distributions by lottery type

Classification

Predict the Lottery Selection

- LDA
- QDA
- Logistic
- Random Forest Adaptive Boosting
- Bagging
- Gradient Boosting

Regression

Predict the Draft Pick Order

- Linear
- Polynomial
- Random Forest
- Lasso
- Ridge

Dataset & Exploratory Analysis



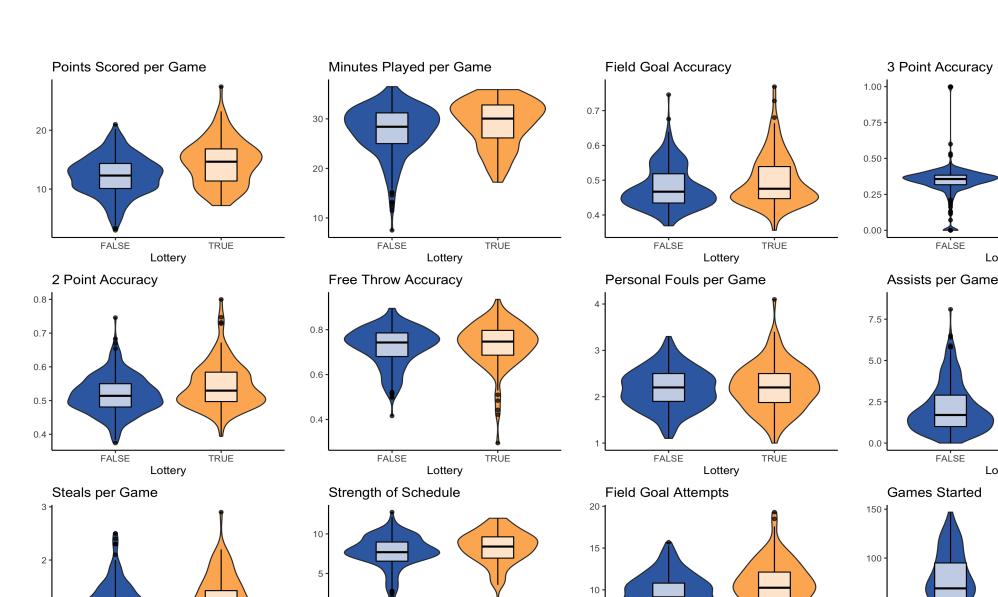
Pk		School	STL	BLK	TOV	PF	sos	name	lottery
	1	Kentucky	1.8	0.5	4	1.9	6.82	john wall	TRUE
	2	Ohio State	1.6	0.7	3.5	2.7	7.86	evan turner	TRUE
	3	Georgia Tech	0.9	2.1	2.5	2.6	9.02	derrick favors	TRUE
	5	Kentucky	1	1.8	2.1	3.2	6.82	demarcus cousins	TRUE
	7	Georgetown	1.5	1.5	2.9	2.5	9.26	greg monroe	TRUE
	9	Butler	1.3	0.9	2	2.1	3.57	gordon hayward	TRUE
	10	Fresno State	2	0.9	2.7	2.8	1.13	paul george	TRUE
	11	Kansas	0.5	2.3	1.2	2.1	8.12	cole aldrich	TRUE
	12	Kansas	1.5	0.5	1.9	1.8	8.25	xavier henry	TRUE
	13	UNC	0.4	2.1	1.4	1.9	9.09	ed davis	TRUE
	14	Kentucky	0.7	1.6	1.6	2.1	6.24	patrick patterson	TRUE
	15	VCU	0.6	2.7	1.5	3.1	-0.69	larry sanders	FALSE
	16	Nevada	0.9	0.8	2.1	2.2	1.98	luke babbitt	FALSE
	18	Kentucky	1.4	0.3	3	2.2	6.82	eric bledsoe	FALSE
	19	Texas	1.3	0.5	1.5	2.4	7.38	avery bradley	FALSE
	21	Iowa State	0.5	1	2	2	7.34	craig brackins	FALSE

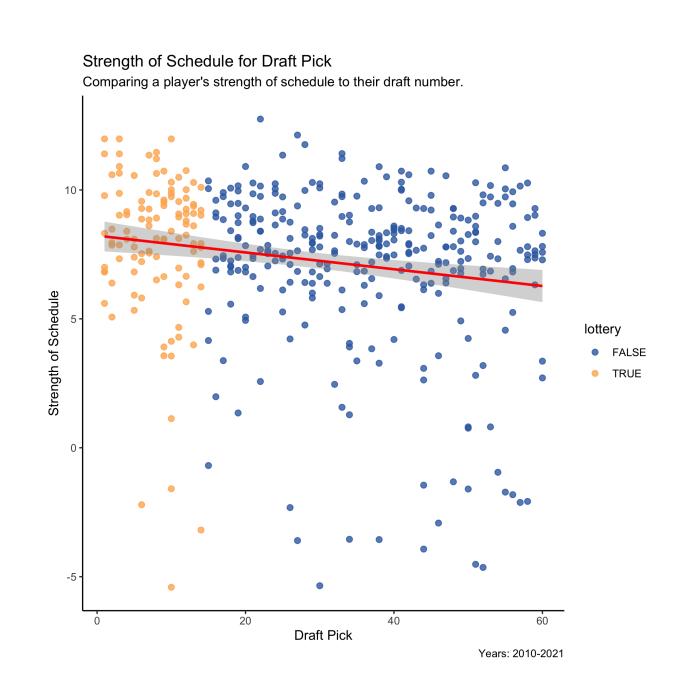
- Historical college basketball statistics scraped from Basketball Reference
- Years: 2010-2021
- Dimensions: 415 Rows,

29 Columns

- College basketball statistics scraped from Basketball Reference
- Years: 2021-2022
- Dimensions: 71 Rows,

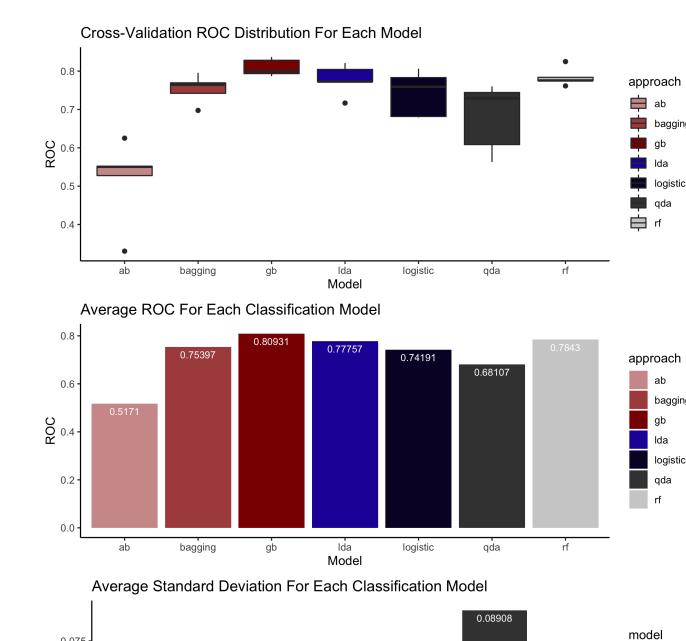
26 Columns

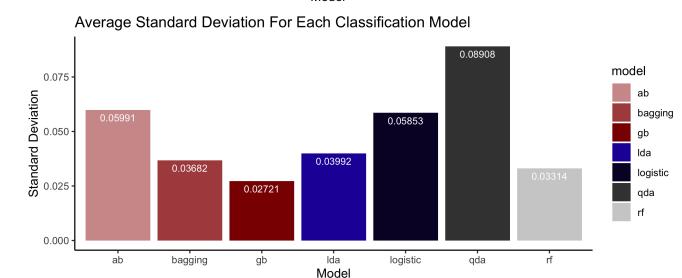




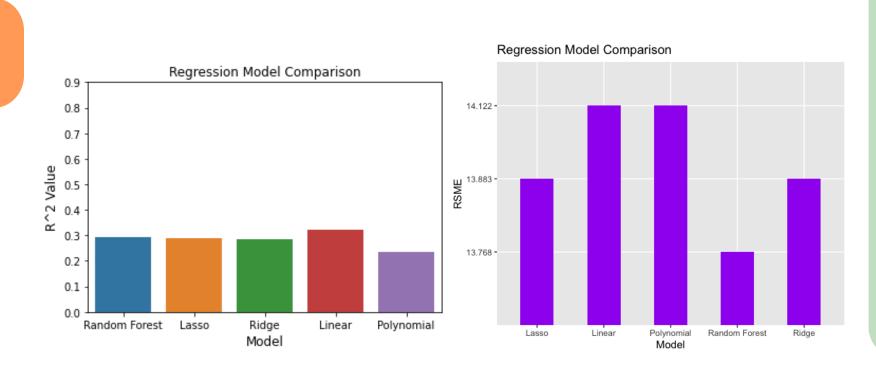
Results

Classification Model Comparison:

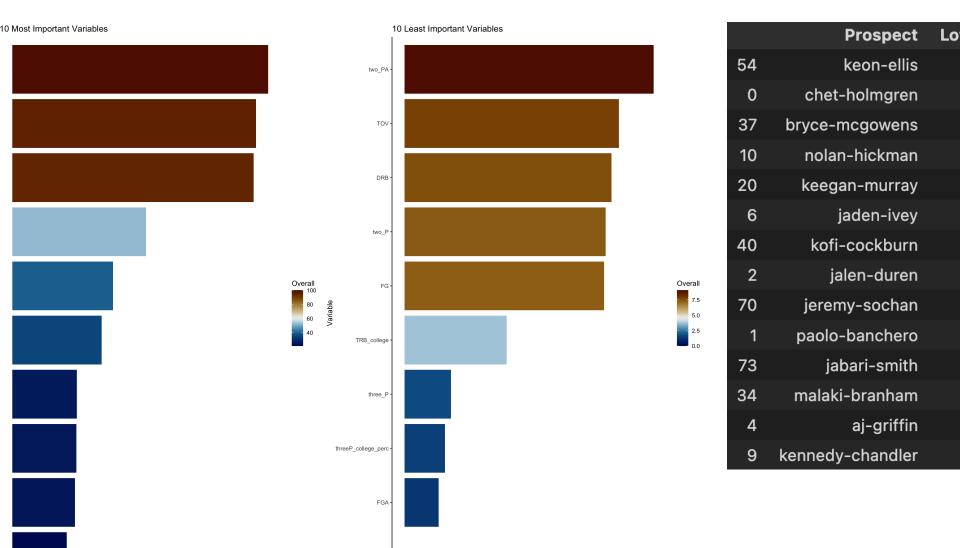




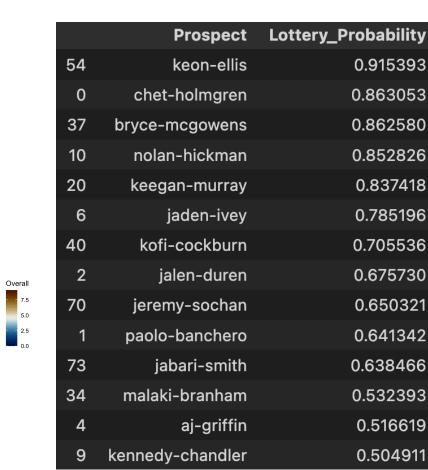
Regression Model Comparison:



Classification Feature Importance:



Lottery Prediction:



Statistical Analysis Summary:

Regression

Lottery, name, and school removed

- Missing values removed Regression models run with all variables
- Adjusted R2 and RMSE compared Optimal model selected Best subset selection to determine most significant
- Interaction terms added to improve accuracy

variables

Classification

Lottery, name, and school removed Missing values removed Lottery converted to factor

- 5-fold cross validation because of data size Random grid search Optimal model determined
- via accuracy and sd comparison
- Features extracted Prospects predicted

Conclusions



Future Directions

- Basketball statistics can accurately predict whether a player will be in the lottery
- 2. Games started, two point percentage, number of games are most important in lottery prediction
- Basketball statistics cannot accurately predict draft pick order
- 1. Assess NBA performance post-draft
- 2. Collect additional demographic information
- 3. Evaluate prospect predictions to actuals
- 4. Segment analysis by position

Literature Cited

- 1. Wikimedia Foundation. (2022, April 1). NBA draft. Wikipedia. Retrieved April 24, 2022, from https://en.wikipedia.org/wiki/NBA_draft#:~:text=The%20NBA%20draft%20is%20an,also%20eligible%20to%20be%20drafted.
- 2. Basketball statistics and history. Basketball. (n.d.). Retrieved April 22, 2022, from https://www.basketball-reference.com/