NBA Draft Prediction

Leveraging college basketball statistics to predict the NBA draft pick

Authors: Kajal Tiwary, Austin Eaton, Jackson Piccione, Abby Fremaux

GEORGETOWN UNIVERSITY

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
Classification — Predict the Lottery Selection
Regression — Predict the Draft Pick Order

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

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

- Linear
- Polynomial
- Random Forest
- Lasso
- Ridge

Dataset

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

MODELING
- Historical college basketball statistics scraped from Basketball Reference
- Years: 2010-2021
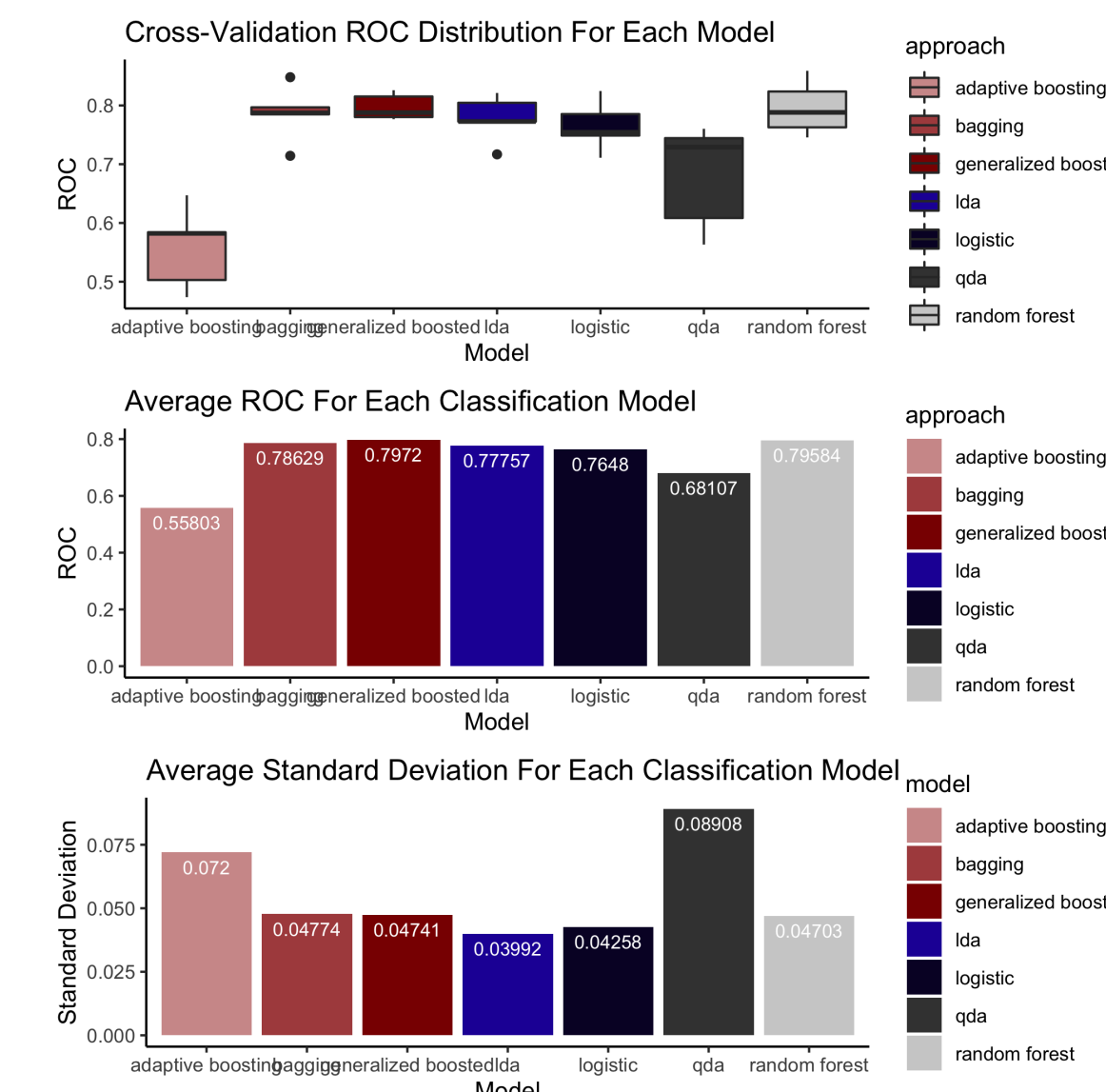- Dimensions: 415 Rows, 29 Columns

PROSPECT
- College basketball statistics scraped from Basketball Reference
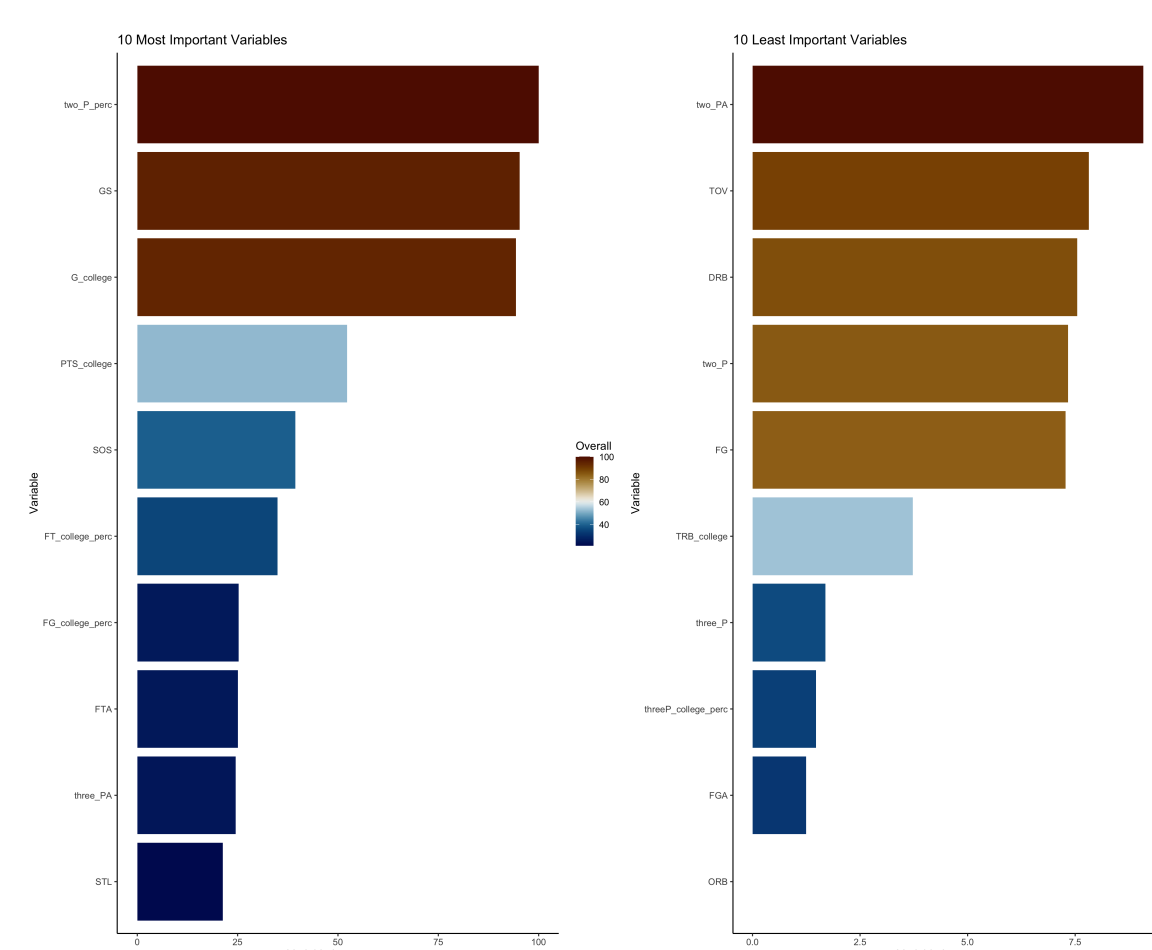- Years: 2021-2022
- Dimensions: 71 Rows, 26 Columns

Violin plots: Points Scored per Game, Minutes Played per Game, Field Goal Accuracy, 3 Point Accuracy, 2 Point Accuracy, Free Throw Accuracy, Personal Fouls per Game, Assists per Game, Steals per Game, Strength of Schedule, Field Goal Attempts, Games Started

Where are most draft picks coming from? Top 20 universities that have contributed players to the NBA draft since 2010. Kentucky leads the pack! Years: 2010-2021

Results

Classification Model Comparison:

Cross-Validation ROC Distribution For Each Model

Average ROC For Each Classification Model

Average Standard Deviation For Each Classification Model

Classification Interpretation:

10 Most Important Variables / 10 Least Important Variables

Prediction:
	Prospect	Lottery_Probability
54	keon-ellis	0.915393
0	chet-holmgren	0.863053
37	bryce-mcgowens	0.862580
10	nolan-hickman	0.852826
20	keegan-murray	0.837418
6	jaden-ivey	0.785196
40	kofi-cockburn	0.705536
2	jalen-duren	0.675730
70	jeremy-sochan	0.650321
1	paolo-banchero	0.641342
73	jabari-smith	0.638466
34	malaki-branham	0.532393
4	aj-griffin	0.516619
	kennedy-chandler	0.504911

Regression Model Comparison:

Regression Model Comparison (R^2 Value and RMSE)

Statistical Analysis Summary:

Regression
- Lottery, name, and school removed to avoid multicollinearity
- Missing values removed
- Regression models run with all variables
- Adjusted R2 and RMSE compared
- Optimal model selected
- Best subset selection to determine most important variables
- Interaction terms added to improve accuracy

Classifcation
- Lottery, name, and school removed to avoid multicollinearity
- Missing values removed
- Lottery converted to factor
- 5-fold cross validation because of data size
- Random grid search
- Hyperparameter tuning
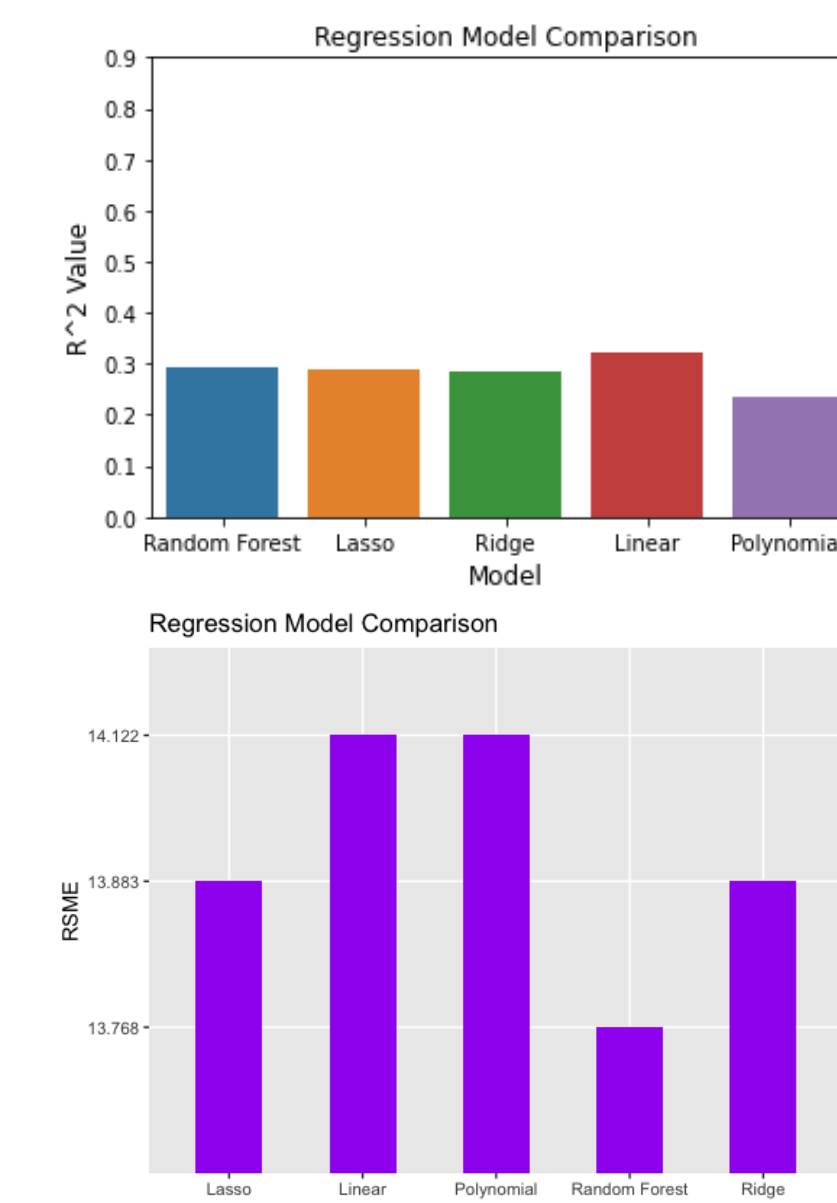
Conclusions

1. 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

3. Basketball statistics cannot accurately predict draft pick order

Limitations and Future Considerations

1. Assess NBA performance post-draft

2. Collect additional demographic information

3. Evaluate prospect predictions to actuals

4. Segment analysis by position

Literature Cited

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.

Basketball statistics and history. Basketball. (n.d.). Retrieved April 22, 2022, from https://www.basketball-reference.com/