

College Football Recruiting Rankings and On-Field Performance



...

Dr. Ross Gosky, Statistics, Appalachian State University
Sydney Singleton, Mathematics, Appalachian State University

Background

- Interest in college football recruiting and team performance has increased
- Several well known websites rank each team's recruits every season
- Questions we had:
 - Are these recruiting rankings predictive of team performance? Do teams that recruit better tend to perform better on the field?
 - If recruiting is predictive of team performance, which components of its measurements are statistically significant?
 - Other questions: Does last season's team success have carryover effects? How much do a team's returning starters matter? Does coaching experience matter?

Project Background

- For years 2007 - 2017, we collected data about college football teams in the largest conferences (AAC, ACC, Big 10, Big 12, Pac 12, SEC, and Notre Dame)
- On field team performance was measured via a team's Sagarin Score (www.sagarin.com)
- Recruiting rankings of each team in each season obtained from Rivals.com
 - Star ranking system
- Other team attributes were gathered from Lindy's College Football preview magazine (www.lindyssports.com/college-football/)

Modeling Team Performance

We used and compared multiple regression models to predict a team's standardized (Z) Sagarin score at the end of the season.

Potential Predictor Variables:

Recruiting Rankings:

Freshman Class

Sophomore Class

Junior Class

Senior Class

Redshirt Senior Class

Recent Team Performance:

Standardized Sagarin Score
last year

Standardized Sagarin Score
two years prior

Bowl game participation and
win in previous year

Team Characteristics:

Returning Starters (offense)

Returning Starters (defense)

Returning Starting QB

Experience of Head Coach

Conference Affiliation

Procedure

Data preparation

- Filled in missing values, took note of some exceptional points
- Formatted for analysis in R/RStudio

- Implemented Forward, backward, stepwise regression
- Used AIC as method of selection
- Produced 3 separate models

Model Selection

Cross-Validation

- Subsetted the chosen season out of the original data.
- Fit each of the three models to the subsetted data

- Predicted the season from each model
- Predicted the top ten in the chosen seasons

Predictions!

Model Selection

Forward
 $R^2=0.52$

- Freshman
class average
stars

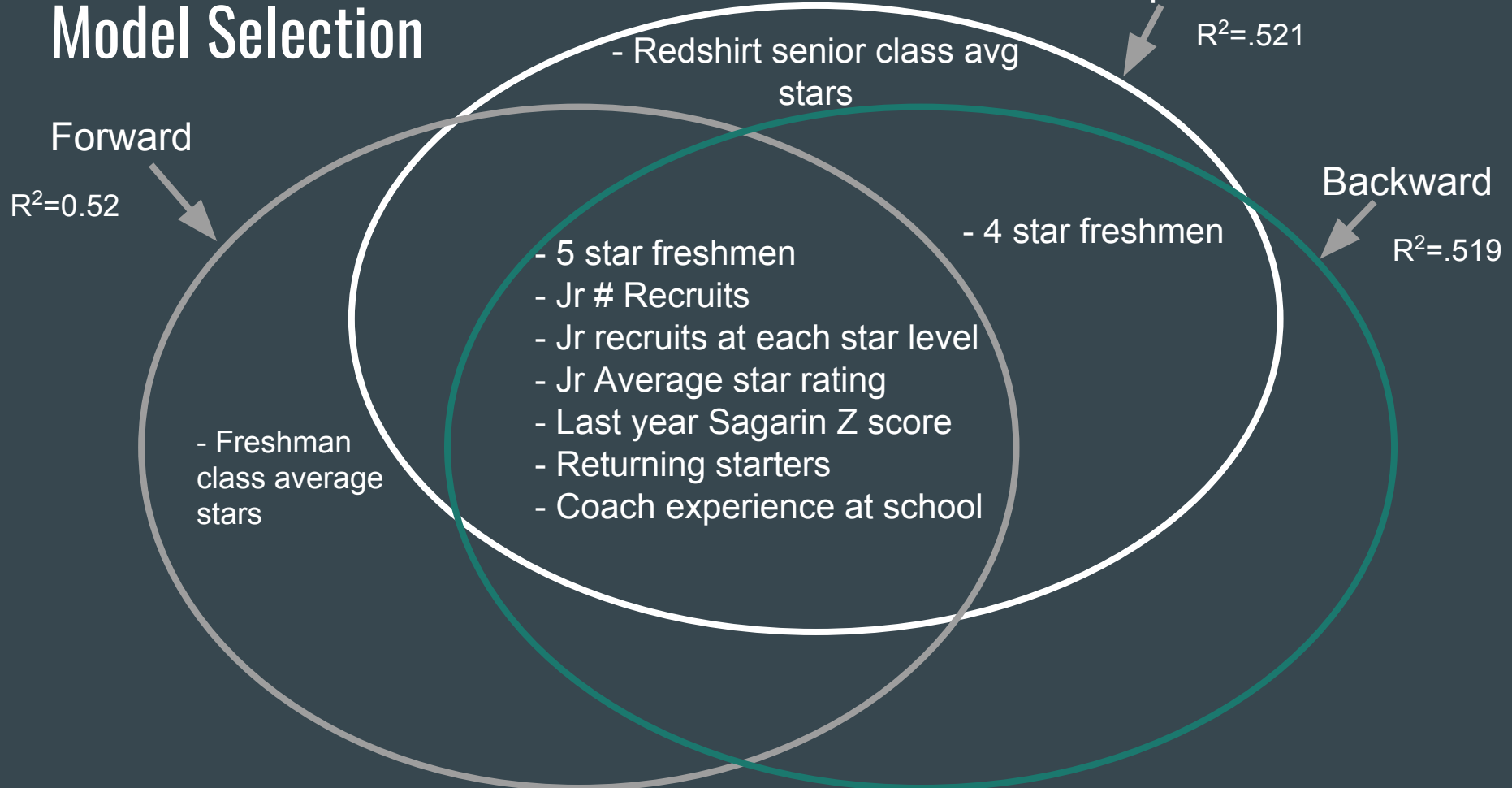
- Redshirt senior class avg
stars

Stepwise
 $R^2=.521$

- 5 star freshmen
- Jr # Recruits
- Jr recruits at each star level
- Jr Average star rating
- Last year Sagarin Z score
- Returning starters
- Coach experience at school

- 4 star freshmen

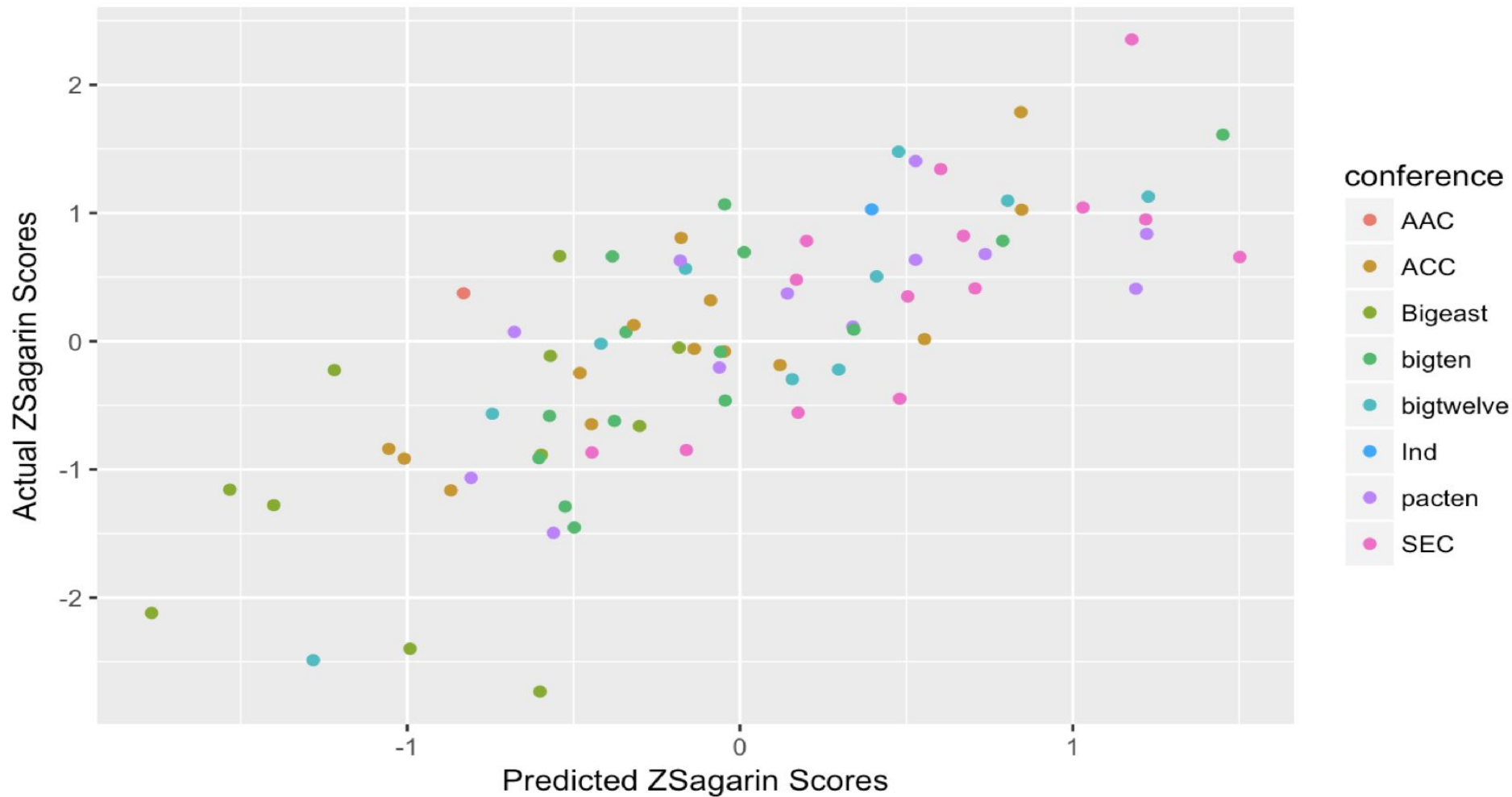
Backward
 $R^2=.519$



Accuracy of the models

- Cross validation: withhold certain observations from the data and use the chosen models to predict their outcomes
- Models with good predictive accuracy should predict these outcomes accurately
- This is a good way to compare models
- Two measures of prediction accuracy:
 - Mean Absolute Percentage Error (MAPE)
 - Mean Squared Prediction Error (MSPE)
- We withheld the 2015 season from the data by design and fit all three models
- Most accurate model was the Stepwise model but all were close
 - MAPE: Forward: 1.80 Backward: 1.81 Stepwise: 1.76
 - MSPE: Forward: .421 Backward: .422 Stepwise: .418

Predicted Vs. Actual ZSagarin Scores



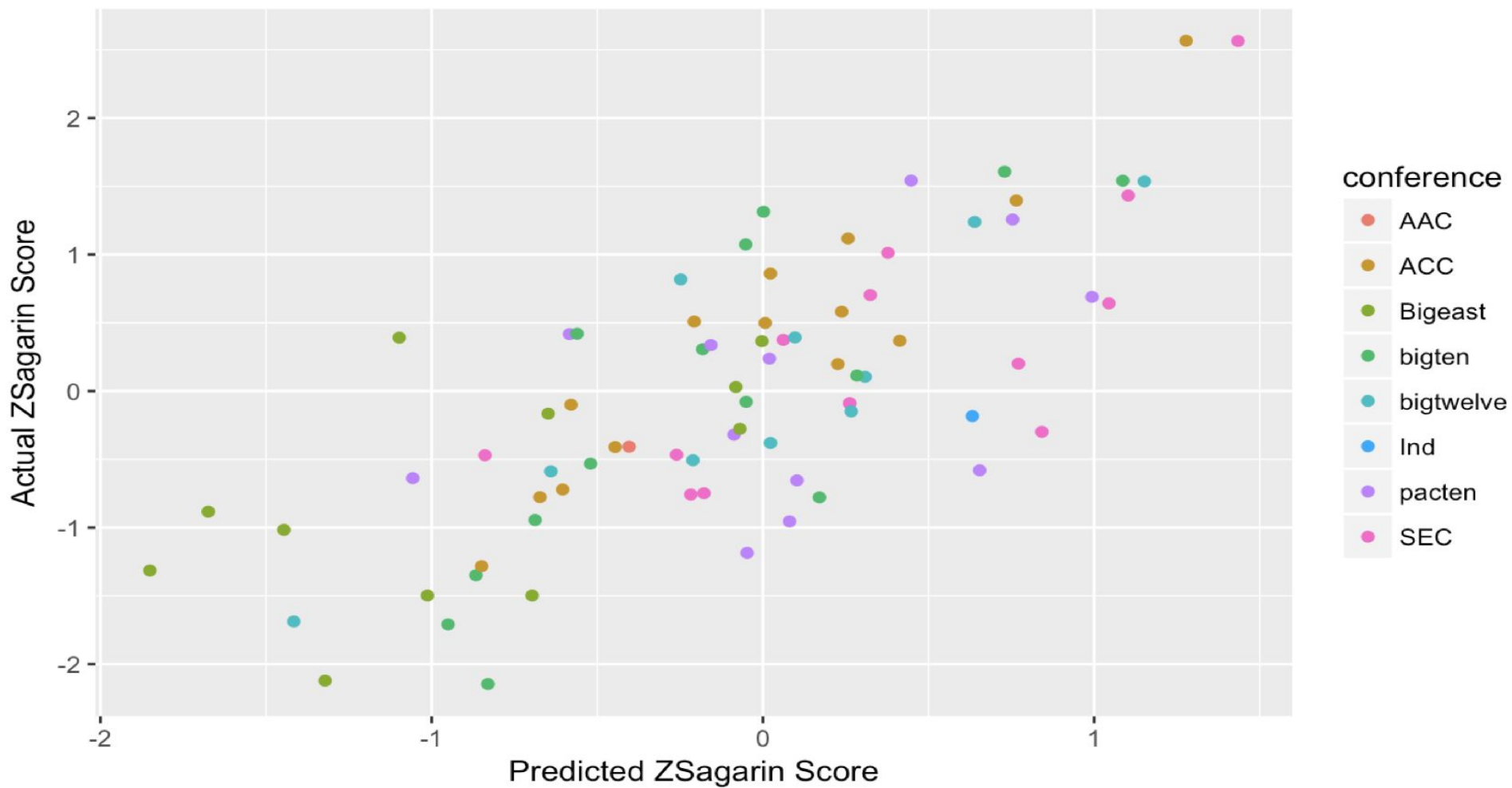
Top Ten Teams: Predicted Vs. Actual 2015

Predicted
Georgia
Ohio State
TCU
Southern Cal
LSU
UCLA
Alabama
Tennessee
FSU
Clemson

Top Ten Teams: Predicted Vs. Actual 2015

Predicted	Actual
Georgia	Alabama
Ohio State	Clemson
TCU	Ohio State
Southern Cal	Oklahoma
LSU	Stanford
UCLA	Ole Miss
Alabama	TCU
Tennessee	Baylor
FSU	Michigan
Clemson	Tennessee

Predictions Vs. Actual for 2016 Season



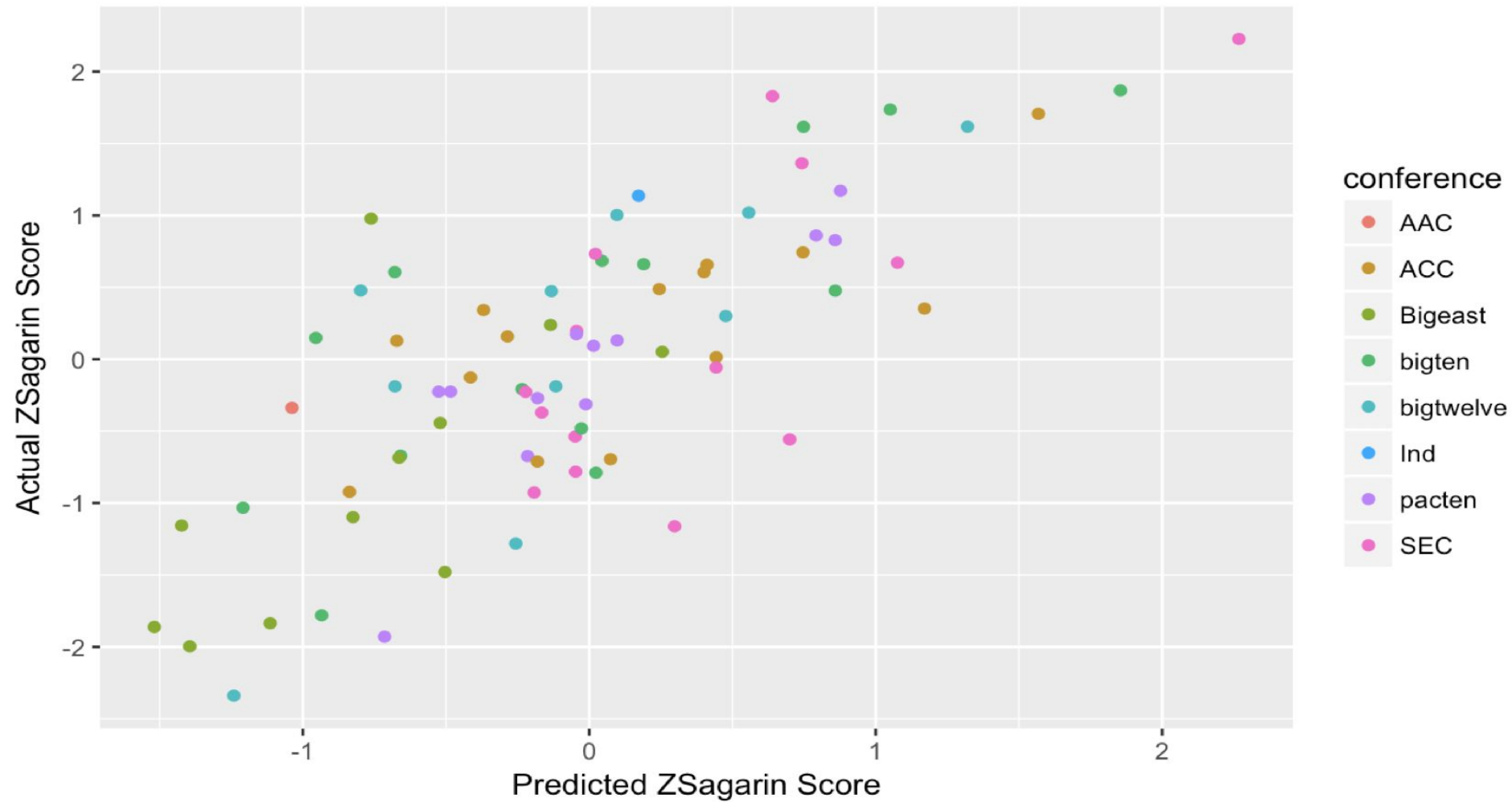
Top Ten Teams: Predicted Vs. Actual 2016

Predicted
Alabama
Clemson
Oklahoma
LSU
Ohio State
Tennessee
Stanford
Ole Miss
Georgia
FSU

Top Ten Teams: Predicted Vs. Actual 2016

Predicted	Actual
Alabama	Clemson
Clemson	Alabama
Oklahoma	Michigan
LSU	Washington
Ohio State	Ohio State
Tennessee	Oklahoma
Stanford	LSU
Ole Miss	FSU
Georgia	Wisconsin
FSU	Southern Cal

Predictions Vs. Actual for 2017 Season



Top Ten Teams: Predicted Vs. Actual 2017

Predicted
Alabama
Ohio State
Clemson
Oklahoma
FSU
LSU
Penn State
Washington
Michigan
Southern Cal

Top Ten Teams: Predicted Vs. Actual 2017

Predicted	Actual
Alabama	Alabama
Ohio State	Ohio State
Clemson	Georgia
Oklahoma	Penn State
FSU	Clemson
LSU	Oklahoma
Penn State	Wisconsin
Washington	Auburn
Michigan	Washington
Southern Cal	Notre Dame

Trivia Question

Who won the college football national championship in 2009?

Trivia Question

Who won the college football national championship in 2009?

University of Florida

- And this is what our model predicted!

Conclusions

- Recruiting does matter!
 - According to our models, if two teams were equally as good last year, but their newest recruiting classes differ, the team with the higher-ranked recruiting class is predicted to do better.
- The junior class is an important part of predicting a team's performance.
- Coach experience matters
- Last year's success is important

Questions?