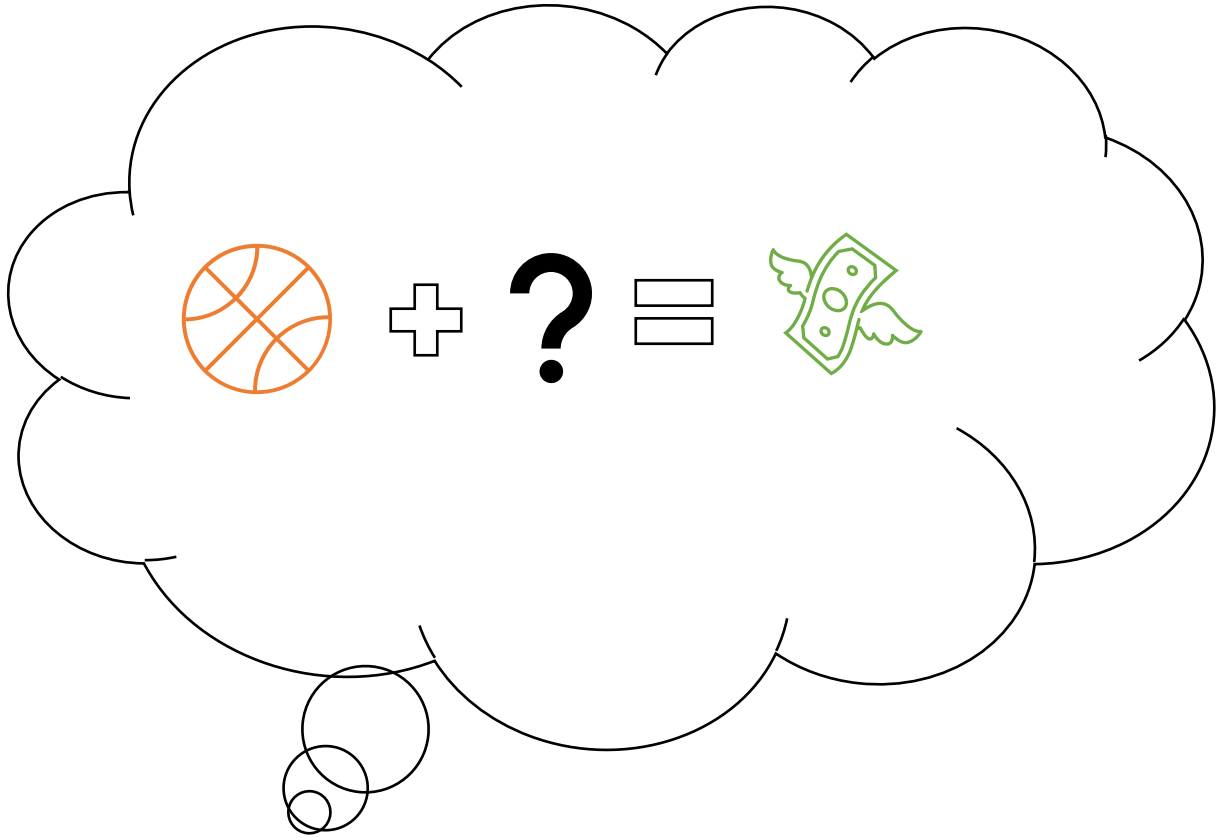
NCAA Tournament Predictions

Emily Edwards, Ben Facio, Alexis Pitzen, Paul Sirena

Overview

- Problem Description
- Data Description
- Modeling
- Evaluation
- Recommendation

Problem Description



Data Description

Target Variable = Qualified

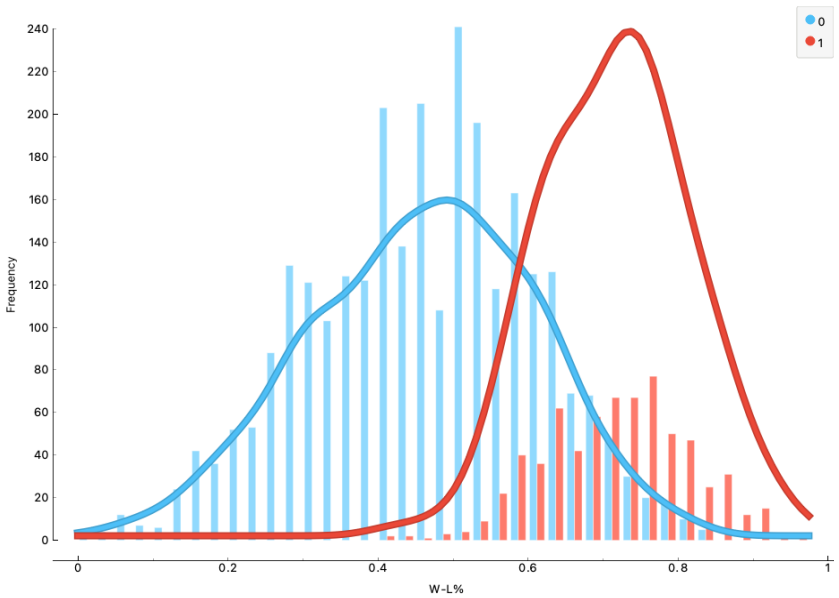
Removed teams that did not play in regular season

Transformations:

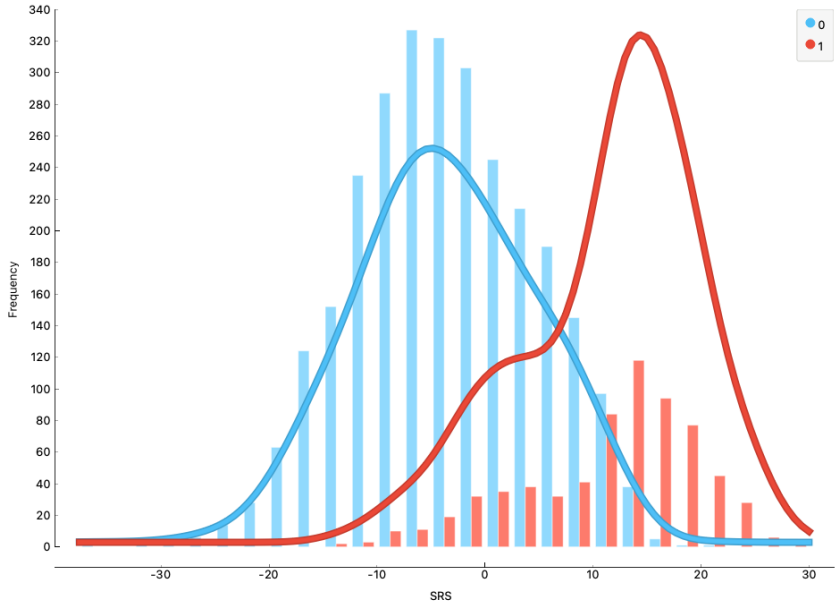
- Converted number of win/losses to percentage
- Normalized SRS, SOS, Pace, PPG, OPPG
- Calculated PPG and OPPG as Total Points/Number of Games
- Binary coded schools in the Power 5 (ACC, Big 10, Big 12, Pac 12, SEC) = 1 and all other Mid Major schools = 0

Feature	Description
Win/loss percent (WL%)	Total wins / total games played
Simple Rating System (SRS)	Average point differential and strength of schedule. Teams ranked above/below average (0)
Strength of Schedule (SOS)	The quality of opponents played
Conference	Power 5 or Mid Major conference
Scored points per game (PPG)	Total points scored in current season
Allowed points per game (OPPG)	Total points scored by opponents in current season
Pace	Estimate of possessions per game
True shooting percent (TS%)	Measure of shooting efficiency incorporating all aspects of shooting
Offensive rebound percentage (ORB%)	Percentage of available offensive rebounds secured by a team
Assist percentage (AST%)	Out of all field goals made, the percentage that were assisted
Steal percentage (STL%)	Percentage of turnovers

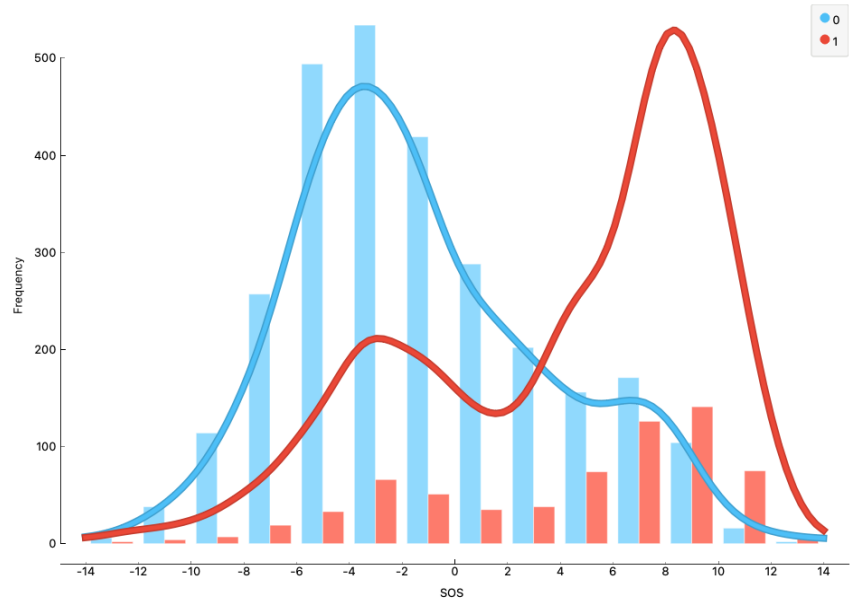
Exploratory Analysis



Win/Loss Percentage



Simple Rating System



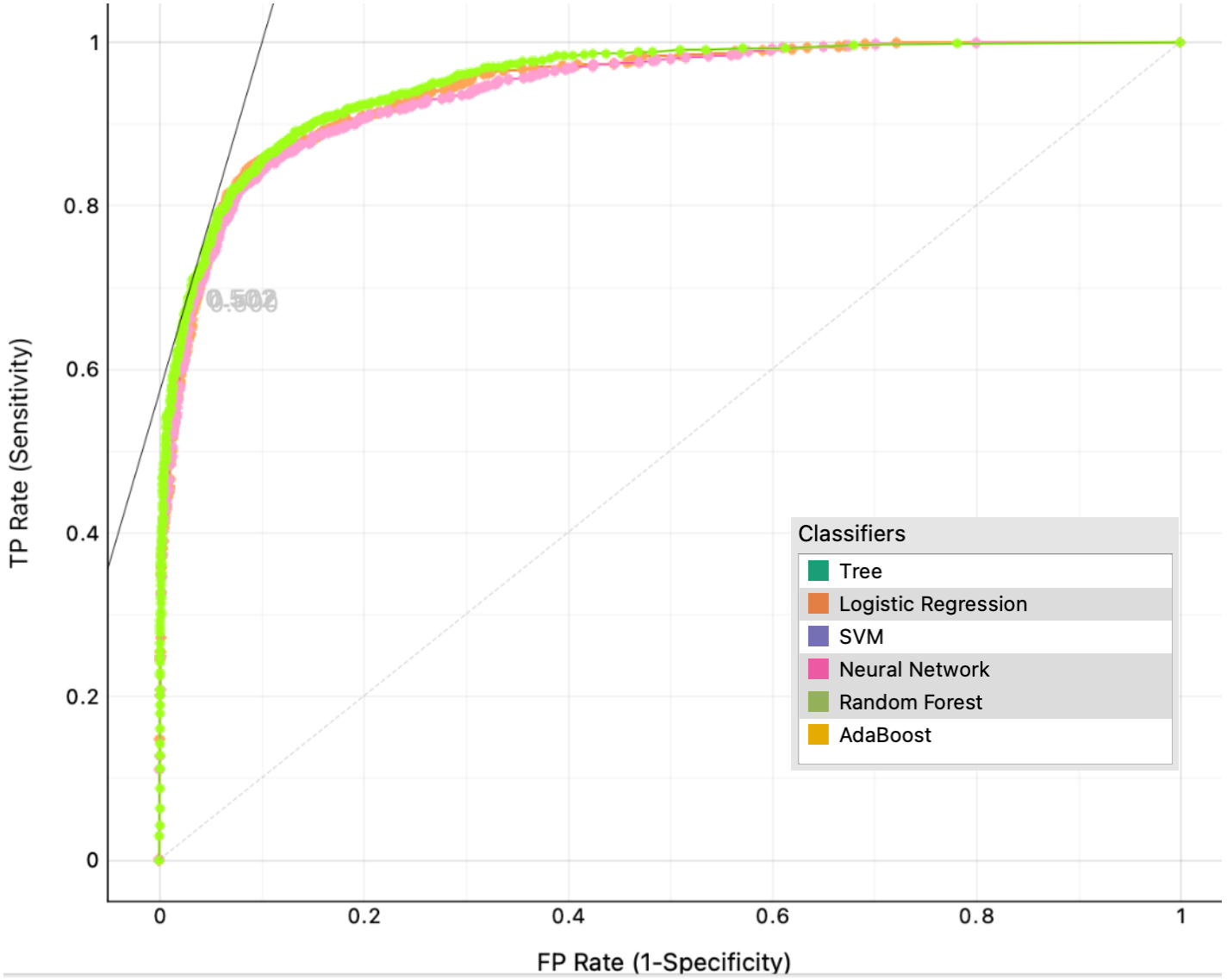
Strength of Schedule

Modeling

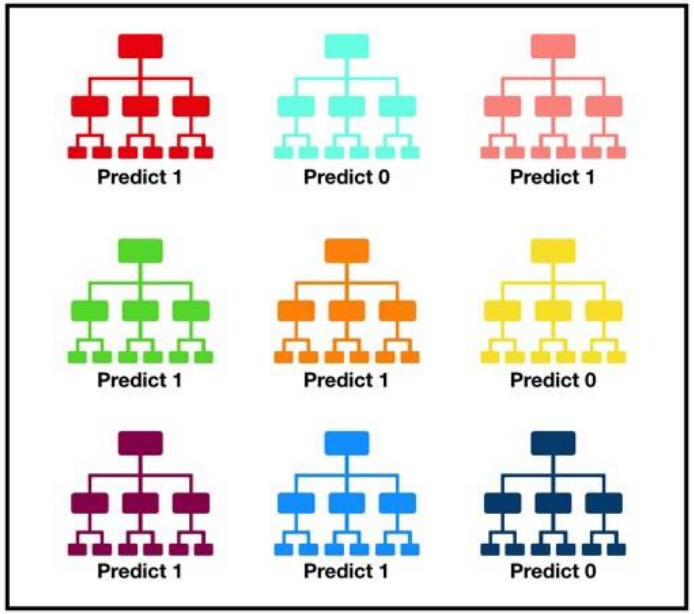
Cross Validated

Target class = Qualified

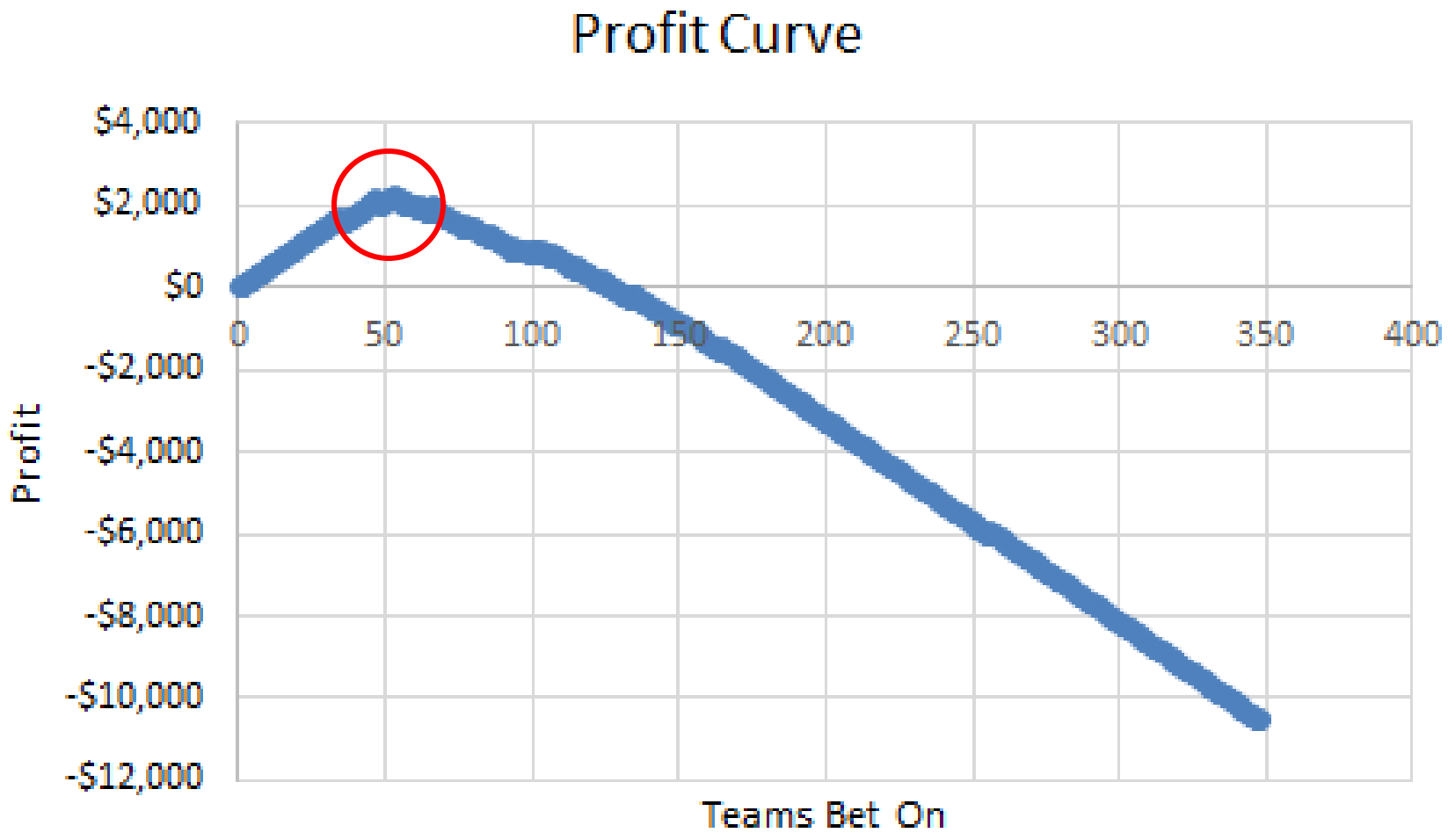
Model	AUC
Random Forest	.951
Logistic Regression	.944
Neural Network	.941
Decision Tree	.920
SVM	.882
Ada Boost	.803



Best Model: Random Forest



Problem | Data | **Model** | Evaluate | Recommend

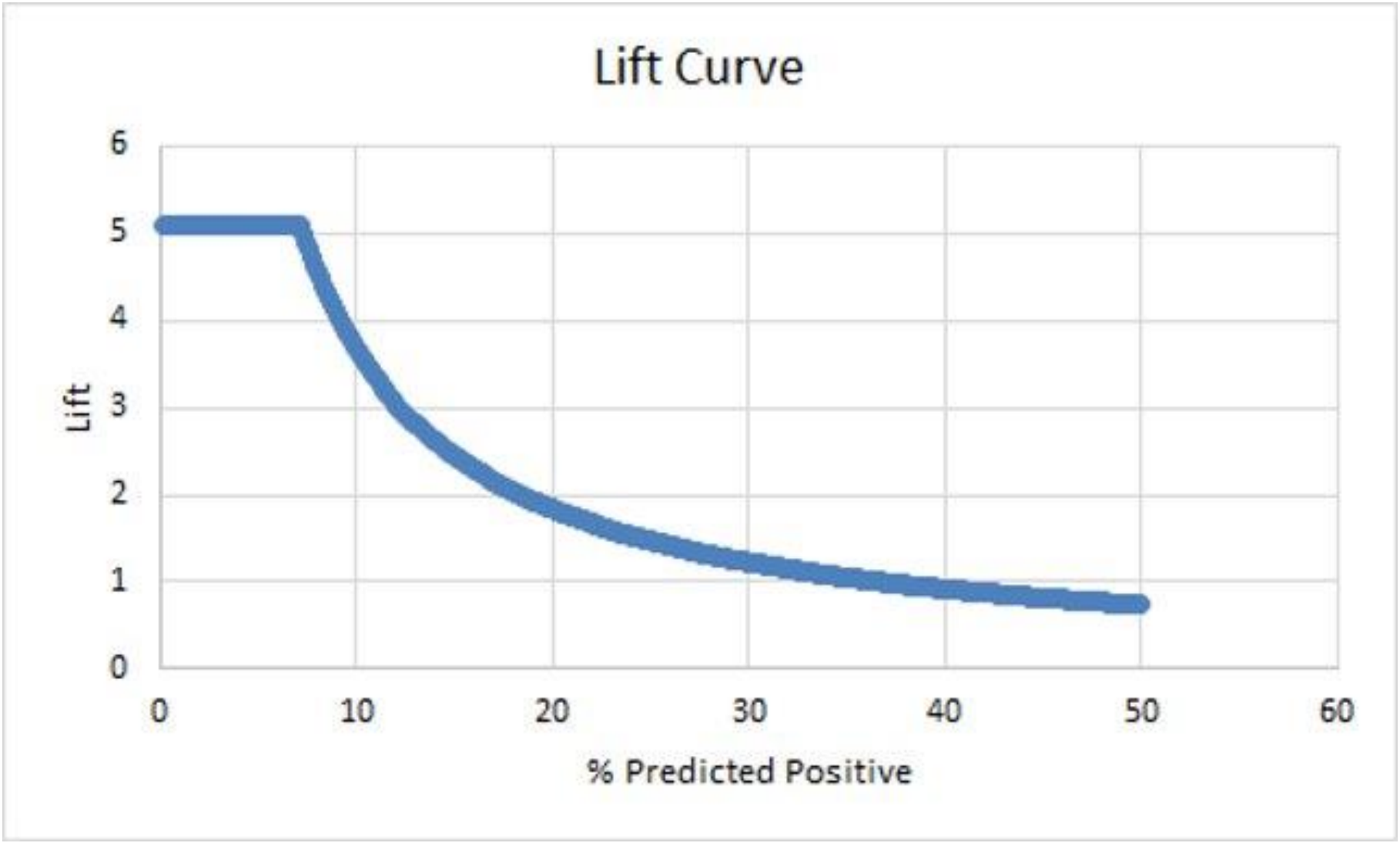


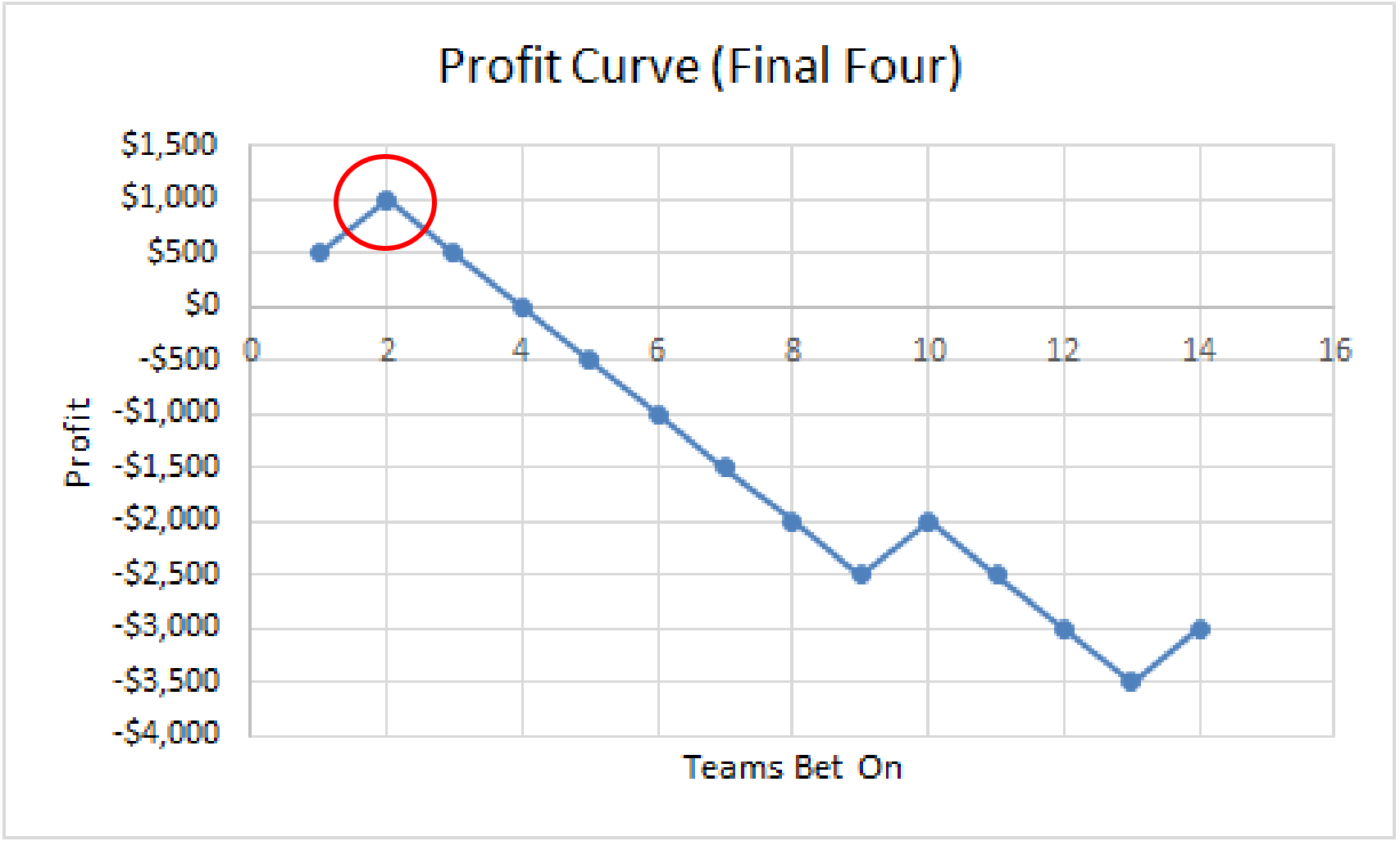
Profit Matrix

	+	-
+	$50	$0
-	($50)	$0

That's a lot of money!

Random Forest vs. Guessing





Final Four Profit Matrix

	+	-
+	$500	$0
-	($500)	$0

2020 NCAA DIVISION I MEN'S BASKETBALL CHAMPIONSHIP BRACKET

NCAA.com

First Round MARCH 19–20 | Second Round MARCH 21–22 | Regional Semifinals MARCH 26–27 | Regional Finals MARCH 28–29 | National Semifinals APRIL 4

FIRST FOUR

DAYTON
MARCH 17–18

Watch On truTV

Grand Canyon	James Madison
Drake	Maryland
Rutgers	Louisville
Siena	Ohio

FINAL FOUR

ATLANTA
APRIL 4 AND 6

NATIONAL CHAMPIONSHIP

APRIL 6

TBS | TBS | TBS

MARCH MADNESS

#MarchMadness

Watch the tournament on these networks or online at NCAA.COM/MARCHMADNESS

DON'T BET ON IT

1	Illinois						Iowa	1
16	GCU/JMU						Siena/Ohio	16
8	BYU						Connecticut	8
9	Texas Tech						Memphis	9
5	North Carolina						Houston	5
12	Drake/Maryland						Rutgers/Louisville	12
4	Oregon						Oklahoma State	4
13	UCSB						Winthrop	13
6	Florida State						Purdue	6
11	Wichita State						Oklahoma	11
3	West Virginia						LSU	3
14	UAB						Bryant	14
7	San Diego State						Michigan	7
10	Virginia Tech						Clemson	10
2	Alabama						USC	2
15	Prairie View						Georgia State	15
1	Colorado						Creighton	1
16	Furman						UMBC	16
8	St Bonaventure						Wisconsin	8
9	Syracuse						Arizona	9
5	Villanova						Oregon State	5
12	Florida						Belmont	12
4	Gonzaga						Virginia	4
13	Weber State						Wright State	13
6	Texas						Tennessee	6
11	Colgate						Georgia Tech	11
3	UCLA						Kansas	3
14	Abilene Christian						Liberty	14
7	Baylor						Loyola (IL)	7
10	VCU						Missouri	10
2	Arkansas						Ohio State	2
15	Norfolk State						South Dakota St	15

1 ⇨ 4

1 ⇦ 5

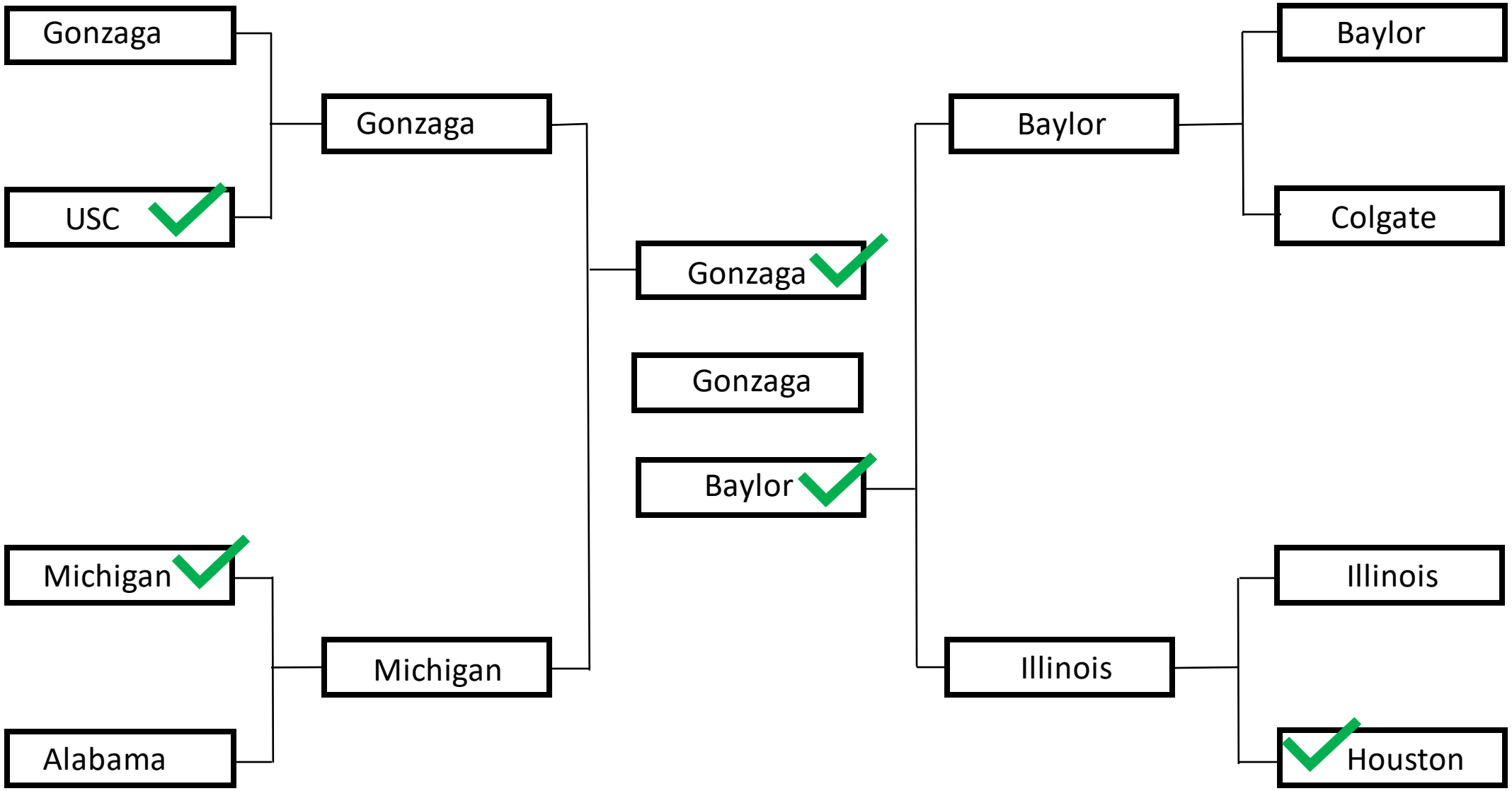
1 ⇨ 7

1 ⇦ 5

In

Out

Bid Stealer:



Problem | Data | Model | **Evaluate** | Recommend

1070 points

91.4 Percentile

Problem | Data | Model | **Evaluate** | Recommend

Model Drawbacks:

Outside Bias

Seeding Secrecy

Abnormal Test Season

Recommend

Bet on top 52 teams to make tournament

Fill out bracket based on Final Four probabilities

Questions?