

WNCAA MARCH MADNESS



Overview



- Business Understanding
- Exploring the Data
- Modeling
- Evaluating Our Models
- Recommendations
- Next Steps



Business Understanding

- NO focus on women's sports
- Gap in the market for sports betting (20% of fans that bet are women)
- Has been an increase, but still 250% more men in online sports gambling
- Growing engagement with women in general = market opportunity

GOAL :

- Use Machine Learning to predict outcome of Women's March Madness, 2023
- Add this as a VIP feature?



**Chance of Correctly
Predicting MM??**

1 in over 9 QUINTILLION
(+9,000,000,000,000,000,000!!!)



Exploring the Data

- Regular Season Detailed Results
 - Winning and Losing team data for every game since 2010
- WNCAA Tourney Detailed Results
 - Winning and Losing team data for each game in every tournament since 2010
- WNCAA Tourney Seeds
 - Seed of each team chosen for the tournament every year since 2010
- 2023 End of Season Team Stats
 - Scrubbed from NCAA, stats for each team in Division I, but was not really used

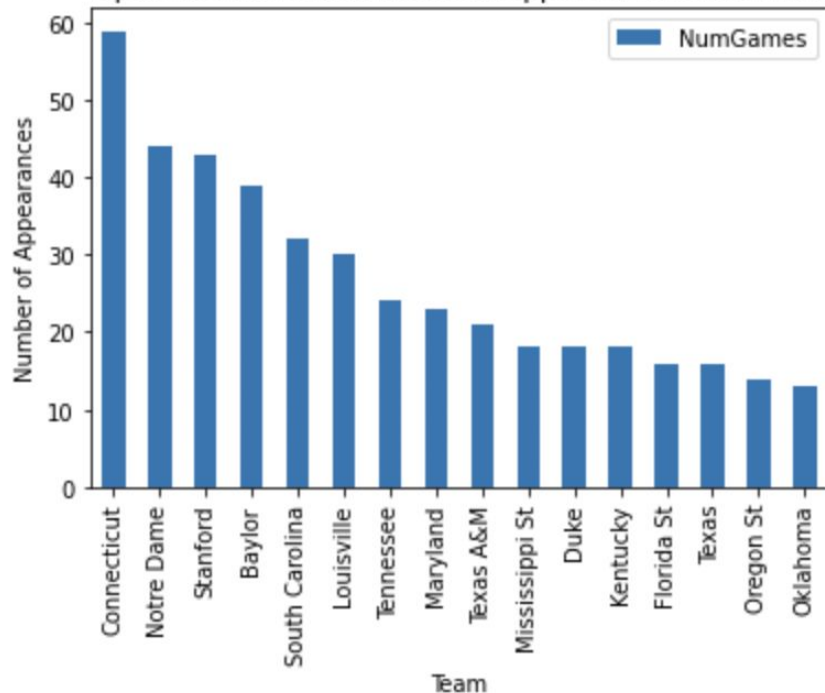


Exploring the Data

	Season	DayNum	TeamName	WScore	TeamName	LScore
0	2021	151	Stanford	66	South Carolina	65
1	2021	151	Arizona	69	Connecticut	59
2	2021	153	Stanford	54	Arizona	53

	Season	DayNum	TeamName	WScore	TeamName	LScore
0	2022	151	South Carolina	72	Louisville	59
1	2022	151	Connecticut	63	Stanford	58
2	2022	153	South Carolina	64	Connecticut	49

Top 16 NCAAW March Madness Appearances Since 2010



Modeling

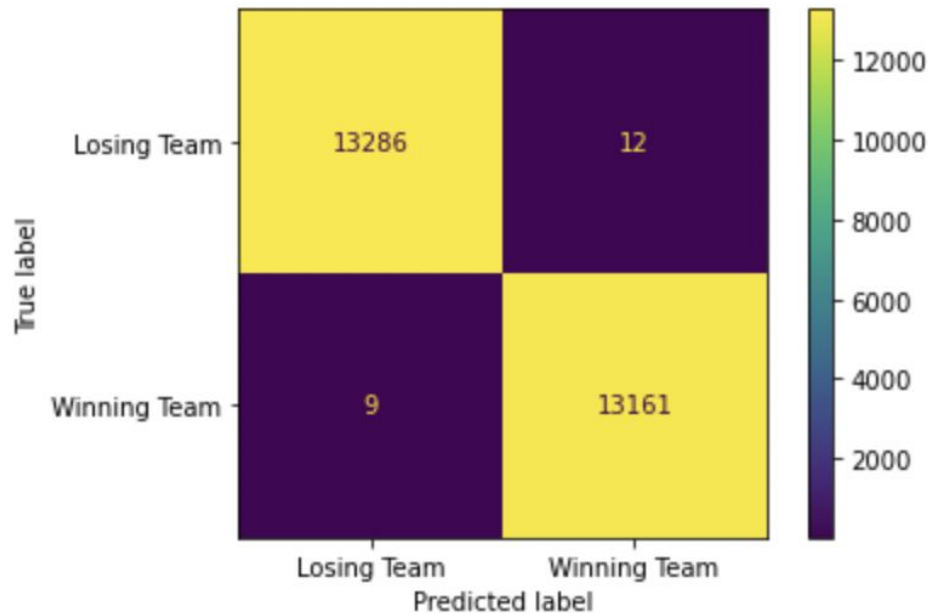


- Decision Trees
- Pipeline with Cross Validation
- XGBoost



Model Evaluation

- Accuracy score:
0.9997733111682031
- Most Important Features
 - Winners Score
 - Loser Score
 - Score Gap
- Most of our instances were correctly classified



Recommendations

1	South Carolina	79
16	Howard	21
Columbia, South Carolina – Fri/Sun		
8	Miami (FL)	78
9	South Florida	66
5	North Carolina	79
12	Stephen F. Austin	66
Tucson, Arizona – Sat/Mon		
4	Arizona	72
13	UNLV	67
6	Georgia	70
11	Dayton	54
Ames, Iowa – Fri/Sun		
3	Iowa State	78
14	UT Arlington	71
7	Colorado	74
10	Creighton	84
Iowa City, Iowa – Fri/Sun		
2	Iowa	98
15	Illinois State	58
1	South Carolina	49
8	Miami (FL)	33
1	South Carolina	69
5	North Carolina	61
5	North Carolina	63
4	Arizona	45
1	South Carolina	80
10	Creighton	50
6	Georgia	44
3	Iowa State	67
3	Iowa State	68
10	Creighton	76
10	Creighton	64
2	Iowa	62



- Focus on Scoring
 - Seemed most important to our model
- Seeding is Important
- Paying attention to “upset” data
 - Player injuries, last few games, etc.

NEXT STEPS

- Digging deeper into other data
 - Upset data
 - Player data such as injuries
 - More specific stats
- Pulling in end of season team data for 2023
- Using model to predict 2022 tournament to test accuracy
- Follow this year's tournament to see how the results hold up!



Thank You!

Ash Heinke

