














]--- March Machine Learning Mania 2023 – Kaggle Competition ---[







What is March Madness?

- One of the most popular sporting events in the United States.
- Every year around March/April
~19-21 days ~3 weeks

- The season-ending tournament for NCAA Division I
- both the men's and women's leagues.



	March Machine Learning Mania 2023 Forecast the 2023 NCAA Basketball Tournaments Featured · 1033 Teams · 2 months ago	\$50,000 ...
	March Machine Learning Mania 2022 - Men's Predict the 2022 College Men's Basketball Tournament Featured · 930 Teams · a year ago	\$25,000 ...
	March Machine Learning Mania 2022 - Women's Predict the 2022 College Women's Basketball Tournament Featured · 651 Teams · a year ago	\$25,000 ...
	March Machine Learning Mania 2021 - NCAAAM - Spread Predict the margin of victory in the 2021 men's tournament Playground · 96 Teams · 2 years ago	Knowledge ...
	March Machine Learning Mania 2021 - NCAAW - Spread Predict the margin of victory in the 2021 women's tournament Playground · 70 Teams · 2 years ago	Knowledge ...
	March Machine Learning Mania 2021 - NCAAW Predict the 2021 NCAAW Basketball Tournament Playground · 451 Teams · 2 years ago	Prizes ...
	March Machine Learning Mania 2021 - NCAAAM Predict the 2021 NCAAAM Basketball Tournament Playground · 707 Teams · 2 years ago	Prizes ...
	March Machine Learning Mania 2017 Predict the 2017 NCAA Basketball Tournament Playground · 441 Teams · 6 years ago	Swag ...
	March Machine Learning Mania 2016 Predict the 2016 NCAA Basketball Tournament Featured · 596 Teams · 7 years ago	\$25,000 ...
	March Machine Learning Mania 2015 Predict the 2015 NCAA Basketball Tournament Featured · 340 Teams · 8 years ago	\$15,000 ...
	March Machine Learning Mania Tip off college basketball by predicting the 2014 NCAA Tournament Featured · 248 Teams · 9 years ago	\$15,000 ...

	Google Cloud & NCAA® March Madness Analytics Uncover the madness of March Madness® Analytics · 0 Teams · 3 years ago	\$25,000 ...
	Google Cloud & NCAA® ML Competition 2020-NCAAM Apply Machine Learning to NCAA® March Madness® Featured · 0 Teams · 3 years ago	Kudos ...
	Google Cloud & NCAA® ML Competition 2020-NCAAW Apply Machine Learning to NCAA® March Madness® Featured · 0 Teams · 3 years ago	Kudos ...
	Google Cloud & NCAA® ML Competition 2019-Men's Apply Machine Learning to NCAA® March Madness® Featured · 862 Teams · 4 years ago	\$25,000 ...
	Google Cloud & NCAA® ML Competition 2019-Women's Apply Machine Learning to NCAA® March Madness® Featured · 497 Teams · 4 years ago	\$25,000 ...
	Google Cloud & NCAA® ML Competition 2018-Women's Apply machine learning to NCAA® March Madness® Featured · 504 Teams · 5 years ago	\$50,000 ...
	Google Cloud & NCAA® ML Competition 2018-Men's Apply Machine Learning to NCAA® March Madness® Featured · 933 Teams · 5 years ago	\$50,000 ...

- There has been a Kaggle competition for March Madness every season for the last 10 seasons
- In 2020 the tourney didn't happen due to covid

The banner features a close-up of an orange basketball with the 'NCAA' logo and the words 'ALL AMERICAN ATTACK' visible. The text 'Featured Prediction Competition' is in the top left. The main title 'March Machine Learning Mania 2023' is in large white font, followed by the subtitle 'Forecast the 2023 NCAA Basketball Tournaments'. The prize money '\$50,000' and 'Prize Money' are on the right. The Kaggle logo and 'Kaggle · 1,033 teams · 2 months ago' are in the bottom left. A navigation bar at the bottom includes links for Overview, Data, Code, Discussion, Leaderboard, Rules, Team, Submissions, and a 'Late Submission' button.

Featured Prediction Competition

March Machine Learning Mania 2023

Forecast the 2023 NCAA Basketball Tournaments

\$50,000
Prize Money

Kaggle · 1,033 teams · 2 months ago

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Team](#) [Submissions](#) [Late Submission](#) ...

- Use the provided data to generate features
- Optionally add some other data
- Train a model
- Predict the probability each possible game's outcome
- Submit before the tourney starts
- Enjoy the games



Used data

Provided:

- (M/W)RegularSeasonDetailedResults
- (M/W)NCAATourneyDetailedResults
- (M/W)NCAATourneySeeds
- (M/W)TeamConferences
- MMasseyOrdinals_thru_Season2023_Day128

Added:

- WRatings
- Mtournament_team_data



Features

Target:
win=0/1

```
Season
T1_TeamID
T1_FG%_home_mean
T1_3P%_home_mean
T1_point_diff_home_mean
T1_win_home_mean
T1_FG%_away_mean
T1_3P%_away_mean
T1_point_diff_away_mean
T1_win_away_mean
T1_Score_mean
T1_opponent_Score_mean
T1_FGM_mean
T1_FGA_mean
T1_FGM3_mean
T1_FGA3_mean
T1_FTM_mean
T1_FTA_mean
T1_OR_mean
T1_DR_mean
T1_Ast_mean
T1_TO_mean
T1_Stl_mean
T1_BlK_mean
T1_PF_mean
T1_opponent_FGM_mean
```

```
T1_opponent_FGM_mean
T1_opponent_FGA_mean
T1_opponent_FGM3_mean
T1_opponent_FGA3_mean
T1_opponent_FTM_mean
T1_opponent_FTA_mean
T1_opponent_OR_mean
T1_opponent_DR_mean
T1_opponent_Ast_mean
T1_opponent_TO_mean
T1_opponent_Stl_mean
T1_opponent_BlK_mean
T1_opponent_PF_mean
T1_FG%_mean
T1_opponent_FG%_mean
T1_3P%_mean
T1_opponent_3P%_mean
```

```
T1_opponent_3P%_mean
T1_point_diff_mean
T1_win_mean
T1_win_ratio_last_month
T1_prev_win_ratio
T1_played_prev_tourney
T1_Seed
T1_OrdinalRank
T1_top1_appearances
T1_top2_appearances
T1_top4_appearances
T1_top8_appearances
T1_top16_appearances
T1_top32_appearances
T1_top64_appearances
T1_top68_appearances
T1_total_tourney_appearances
T1_conf_point_margin
T1_conf_win_margin
seed_diff
```

+ the same attributes for the second team

Modelling

Null values

XGBoost
Classifier

Test data:
last 3 seasons
Training data:
previous seasons

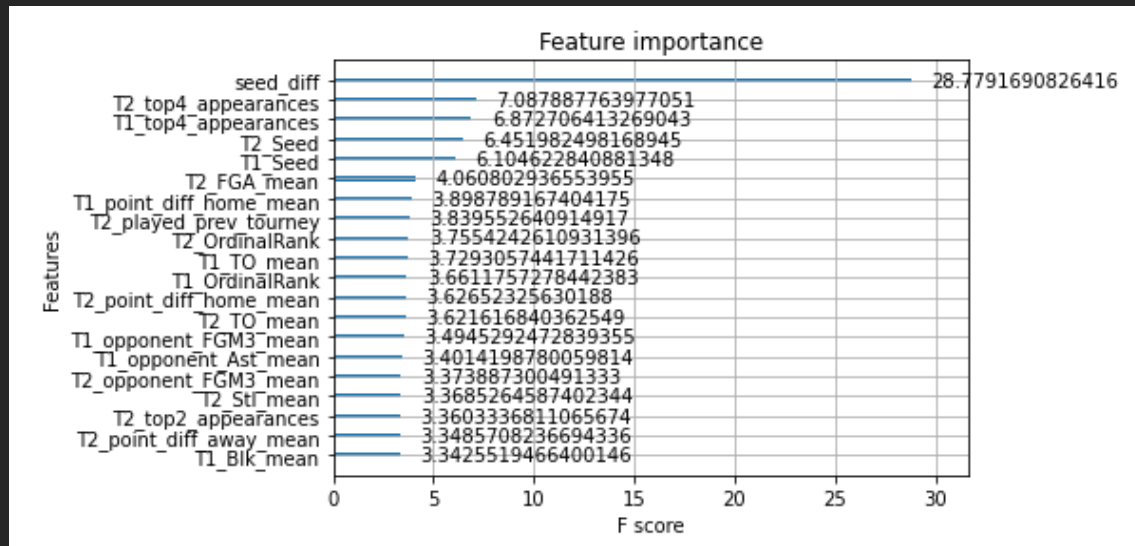
Hyperparameter
tuning with
grid search and
cross-validation

Predicted
probabilities

Brier score
evaluation
metric

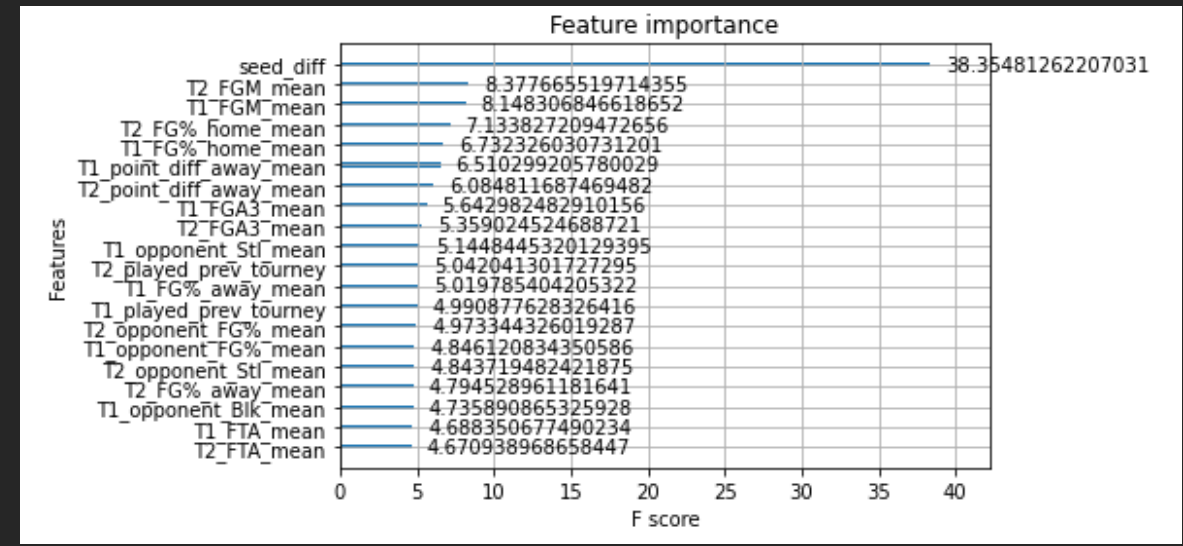
Results

Men



- Brier score: 0.2189
- Accuracy: 0.6685

Women












- Brier score: 0.1602
- Accuracy: 0.7539

Final submission





- Brier score: 0.1936
 - Men: 0.2176
 - Women: 0.1697
- Median Kaggle result: 0.1984
- Best Kaggle result: 0.1737
 - Replacing predicted probabilities with 1s for big favourites
- Accuracy: 0.7143
 - Men: 0.6825
 - Women: 0.7460

ID	Pred
2023_1104_1112	0.723990
2023_1104_1113	0.915681
2023_1104_1116	0.845466
2023_1104_1120	0.845042
2023_1104_1124	0.839218
...	...

Our placing

#	△	Team	Members		Score	Entries	Last	Solution
1	—	RustyB			0.17371	2	2mo	
2	—	Matthias Kullowatz			0.17522	2	2mo	
3	—	tihonby			0.17525	2	2mo	



435	—	Mallorqui			0.19353	2	2mo	
436	—	Tevfik Erkut			0.19355	2	2mo	
437	—	David Mezzetti			0.19370	2	2mo	
438	—	Yuseok Jung			0.19377	2	2mo	



US