



# Does Diana Taurasi Have a Sweet Spot?

Melanie Ackerman  
Metis Classification Project  
2/23/22

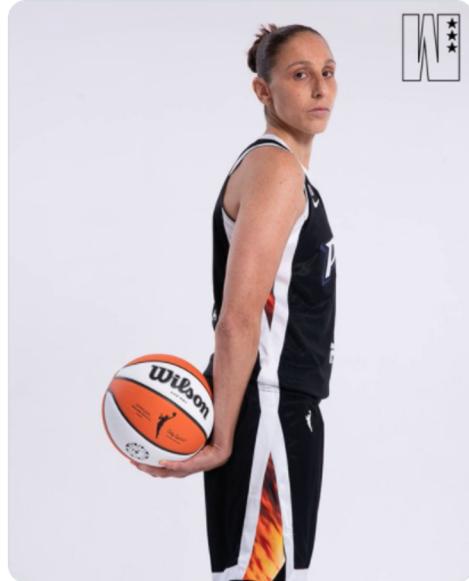
# Who is Diana Taurasi?



WSLAM @wslam · Oct 10

Diana Taurasi has been named the WNBA GOAT.

Salute.



- WNBA superstar on the Phoenix Mercury
- 3x WNBA champion
- 2x WNBA Finals MVP
- 2009 Regular Season MVP
- 1st all-time in career regular season points (9,174)
- 1st all-time in career playoff points (1,397)
- 5th all-time in career regular season assists (2,032)
- 1st all-time in career regular season 3-pointers (1,205)
- And more...

---

# Does it Matter From Where on the Court She Shoots?



DT can shoot from anywhere on the court...

Literally.

...But can we find her sweet spot and set her up to score even more?

---

# The Data

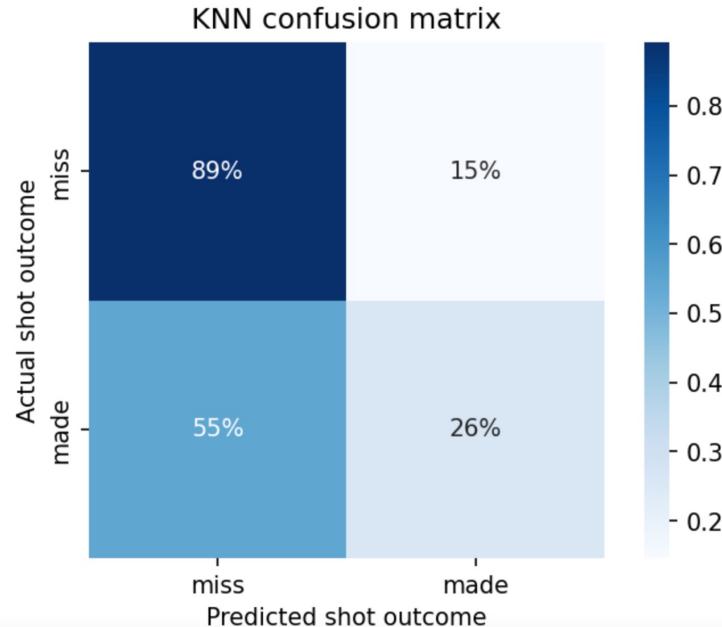
#	Column	Non-Null Count	Dtype
0	GRID_TYPE	8884 non-null	object
1	GAME_ID	8884 non-null	object
2	GAME_EVENT_ID	8884 non-null	int64
3	PLAYER_ID	8884 non-null	int64
4	PLAYER_NAME	8884 non-null	object
5	TEAM_ID	8884 non-null	int64
6	TEAM_NAME	8884 non-null	object
7	PERIOD	8884 non-null	int64
8	MINUTES_REMAINING	8884 non-null	int64
9	SECONDS_REMAINING	8884 non-null	int64
10	EVENT_TYPE	8884 non-null	object
11	ACTION_TYPE	8884 non-null	object
12	SHOT_TYPE	8884 non-null	object
13	SHOT_ZONE_BASIC	8884 non-null	object
14	SHOT_ZONE_AREA	8884 non-null	object
15	SHOT_ZONE_RANGE	8884 non-null	object
16	SHOT_DISTANCE	8884 non-null	int64
17	LOC_X	8884 non-null	int64
18	LOC_Y	8884 non-null	int64
19	SHOT_ATTEMPTED_FLAG	8884 non-null	int64
20	SHOT_MADE_FLAG	8884 non-null	int64
21	GAME_DATE	8884 non-null	object
22	HTM	8884 non-null	object
23	VTM	8884 non-null	object

- Shot-by-shot data scraped from wnba.com during DT's career (17 seasons)
- Use variables related to shot attempts on the court, including location, area, and range.
- Use discrete variables for period of game and season
- Target variable is SHOT\_MADE\_FLAG

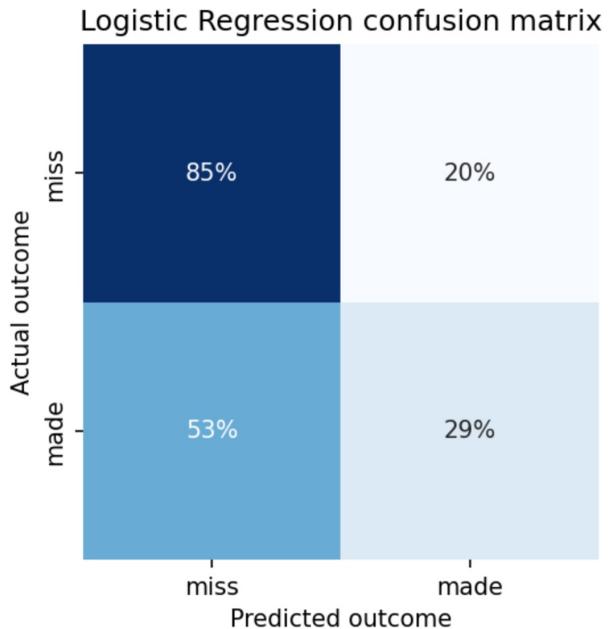
# The Data

- Create dummy variables for the categories included in SHOT\_ZONE\_AREA
  - AREA\_Back Court(BC)
  - AREA\_Center(C)
  - AREA\_Left Side Center(LC)
  - AREA\_Left Side(L)
  - AREA\_Right Side Center(RC)
  - AREA\_Right Side(R)
- and SHOT\_ZONE\_RANGE
  - RANGE\_16-24 ft.
  - RANGE\_24+ ft.
  - RANGE\_8-16 ft.
  - RANGE\_Back Court Shot
  - RANGE\_Less Than 8 ft.

# Can shot coordinates alone predict the outcome of a shot?

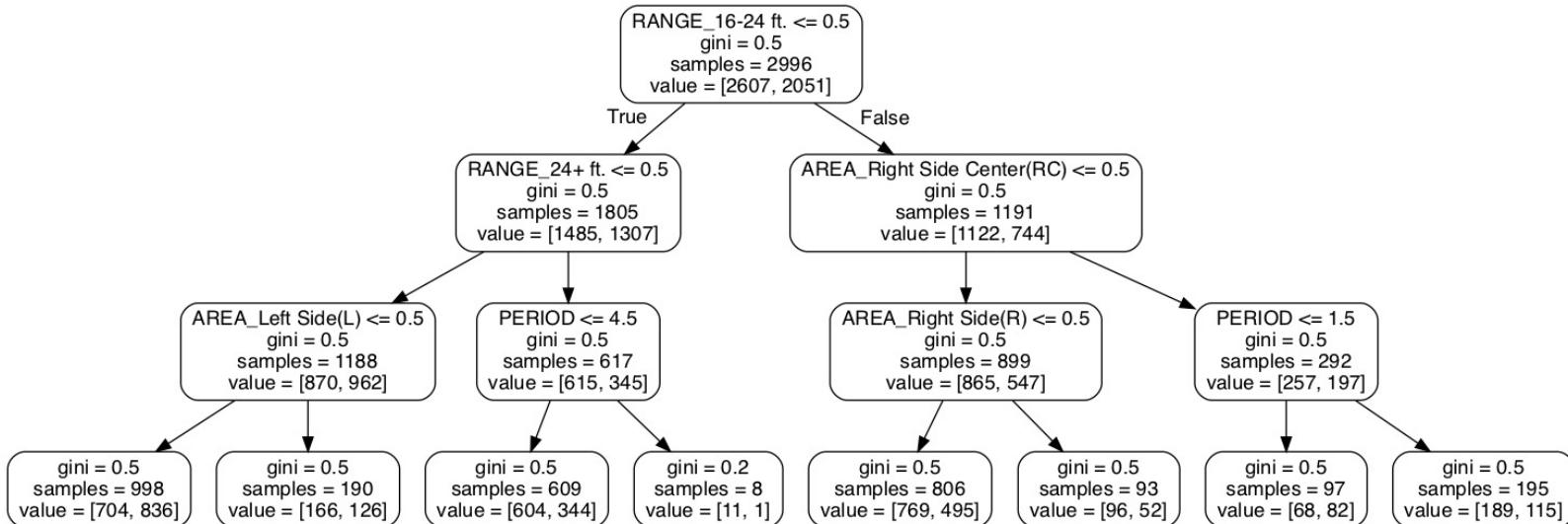


# Logistic Regression



- Baseline model: Logistic Regression
  - Precision: 0.5900621118012422
  - Recall: 0.2865761689291101
  - F1: 0.38578680203045684
- Precision score is priority
- Low interpretability for the data
- Similar performance to KNN

# Random Forest Classifier



Precision: 0.60790273556231

Recall: 0.30721966205837176

F1: 0.40816326530612246

---

# Conclusion: Is DT too good to predict?



- Predicting a sweet spot is challenging using location alone.
  - Play-by-play data would be helpful
  - Or... maybe Diana Taurasi is really just so good at shooting that it doesn't matter much where she shoots from.
-