

## The Hot Hand Phenomenon

**Leon Wang** LJW7736@NYU.EDU  
**Peter Hu** RH3659@NYU.EDU  
**Yuan Li** YL10134@NYU.EDU

## Milestone 3 – Wrap Up

## 1. Methodology

### 1.1 Shot-make Prediction Model using LightGBM

As we discussed in the Milestone 2 report, the hot hand phenomenon is really about whether the Bernoulli random variables for each shot are independent. In controlled environments, such as free throws, those Bernoulli random variables are considered identically distributed. However, in real game situations, the probability of making each shot is different. Such shot-make probabilities must be estimated first, since the testing of independence depends on them.

Pelechrinis and Winston (Pelechrinis and Winston, 2022) use a neural network classifier with 4 hidden layers to predict shot-make probabilities based on various factors, such as distance to the basket, distance of the closest defender, etc. In the paper, Pelechrinis and Winston states that their model achieves an out-of-sample accuracy that is on par with the state-of-the-art shot-make probabilities models.

We build a classification model using the same features as Pelechrinis and Winston. Those features include Distance to the basket, Distance of the closest defender, Player IDs for the shooter and the closest defender, Touch time prior to shooting, and Number of dribbles before shooting. We use Weights and Biases to search for good hyperparameters of LightGBM. Our model use the following:

```
params = {
    'objective': 'binary',
    'metric': 'binary_logloss',
    'boosting_type': 'gbdt',
    'num_leaves': 71,
    'learning_rate': 0.025,
    'min_data_in_leaf': 20,
    'max_depth': 40,
    'feature_fraction': 0.65,
    'verbose': -1,
    'n_estimators': 200
}
```

On the same test data, our model performs slightly better than the model of Pelechrinis and Winston.

## 1.2 Streak Number Test for hot hand

The conditional probability test for hot hand used by GVT (Gilovich and Tversky, 1985) has been the main tool of testing hot hand. Miller and Sanjurjo (Miller and Sanjurjo, 2018) point out the issue of the streak selection bias and suggest using permutation to adjust the bias. Pelechrinis and Winston extend the permutation into simulation, but they still use conditional probability test for hot hand. The conditional probability test is intuitive, but it suffers the streak selection bias, and for higher streak numbers, the noise of the test is high due to the smaller number of samples.

We develop a new test for the hot hand based on the number of streaks. When a player has a hot hand during the game, he tends to have a higher chance of making the next shot after making a streak of shots. As a result, the number of k-streaks ( $k=2, 3, 4$ , etc) that he makes is higher than that of the case of all his shots being independent, even with the fact that the probability of making each shot is not identical. When the shot-make probability is known, we first assume his shots are independent and use simulation to estimate the empirical distribution of k-streaks. Then we can calculate the p-value of his observed number of k-streaks. A very small p-value means the player makes an extraordinarily high number of k-streaks, and it suggests that the player's shots are not independent, i.e., he has a hot hand during that game.

For example, suppose a player plays 100 games for 2 seasons. For each game, he makes 15 shot attempts. Suppose his shot-make (field goal) probability is 45%. Then we can use simulation to find the distribution of the total number of 4-streaks he will make, assuming that each shot is independent.

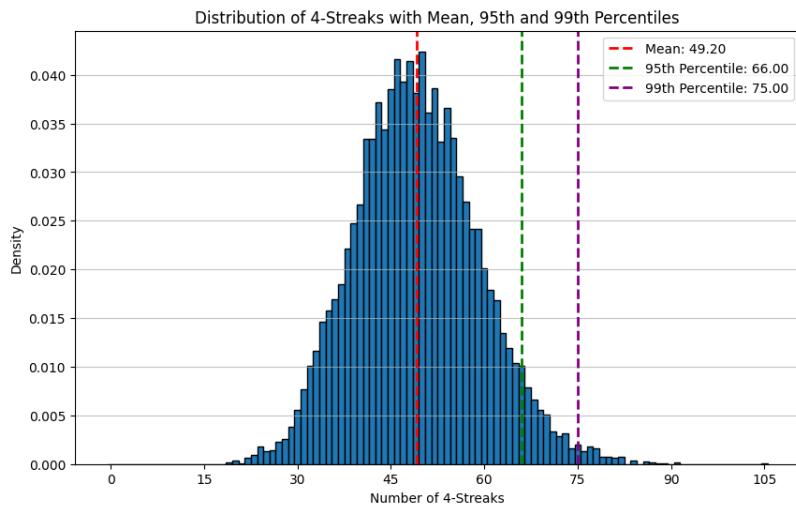


Figure 1: Example of Streak Number Distribution

If the outcomes of each shot of the player are independent, then on average, he makes 49.2 4-streaks in 100 games and the 99% percentile is 75 4-streaks. If he actually makes more than 75 4-streaks, then we can say that, with 99% confidence, we can reject the null hypothesis that his shots are independent.

In our hot hand test, we first use the shot-make prediction model to predict the probability of success for each shot a player makes during a game. Then we perform streak number tests using the actual shooting record of the player.

## 2. Results and Analyses

### 2.1 Shot-make probability prediction models

In the paper of Pelechrinis and Winston, they have two shot-make probability models trained on the data of two seasons (2013 and 2014). Then they test the 2013 trained model on 2014 data and the 2014 trained model on 2013 data. In this way, they avoid data leakage while at the same time maximizing the use of data.

We run the neural network model of Pelechrinis and Winston and find the accuracy at 64.74% and AUROC = 69%. Below is the ROC curve of their model.

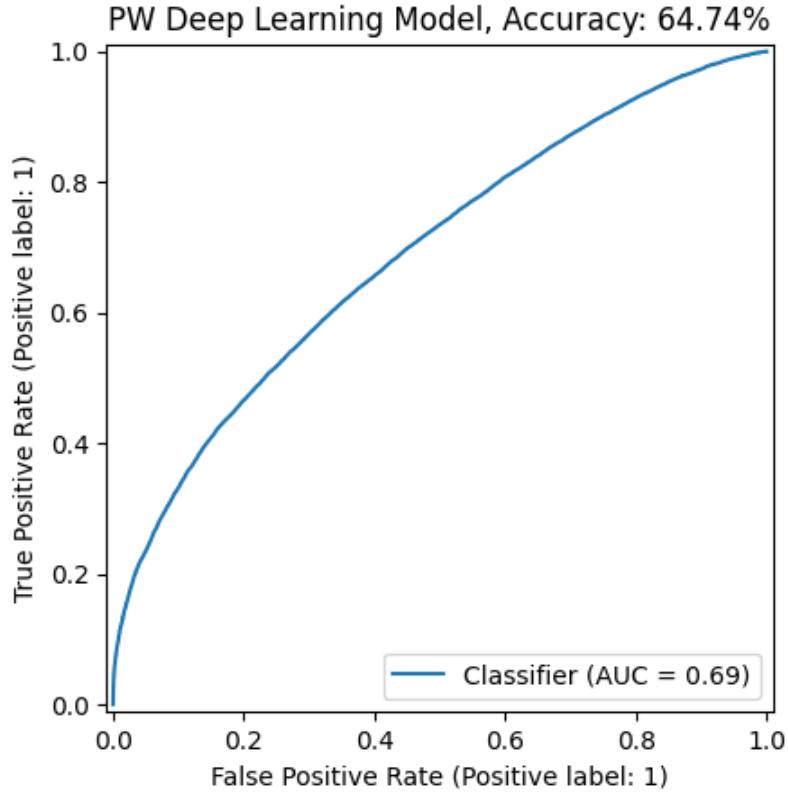


Figure 2: PW Neural Network Model, Accuracy 64.74%

We train and test our LightGBM model based on the same procedure and data as in Pelechrinis and Winston and find the accuracy at 65.66% and AUROC = 70%. Below is the ROC curve of our model.

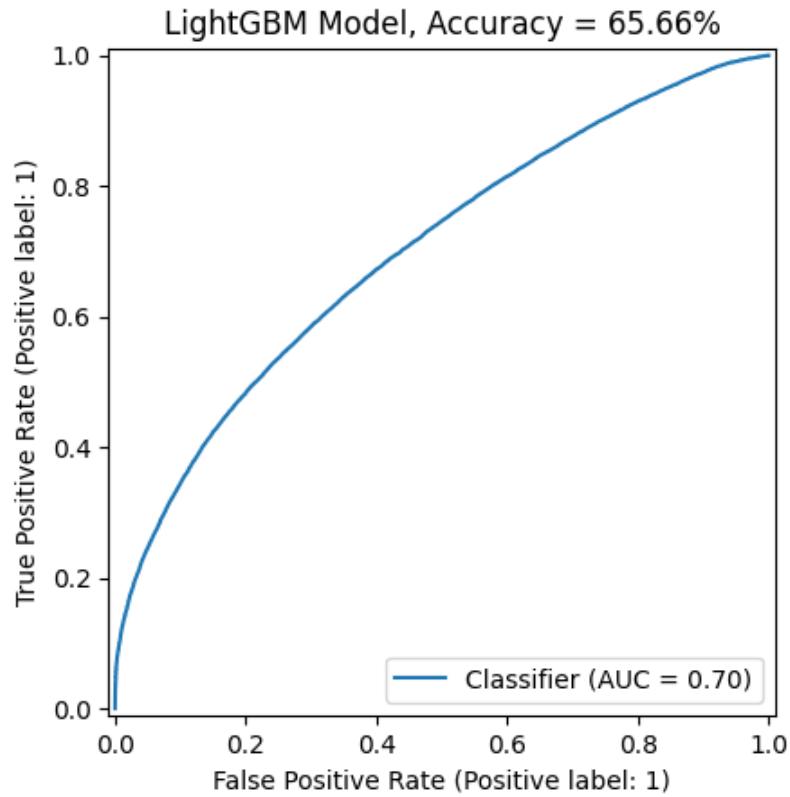


Figure 3: LightGBM Model, Accuracy 65.66%

In the process of finding the hyperparameters for the LightGBM model to maximize the accuracy, we use Weights and Biases random search. Here is a graph of the sweep configuration.

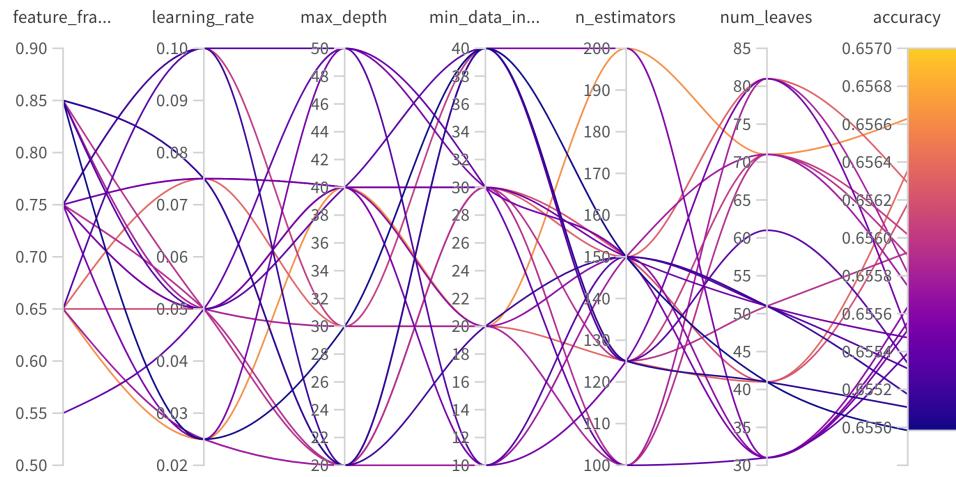


Figure 4: Weights and Biases Random Search for LightGBM hyperparameters

The following are the optimal hyperparameters we have found, which achieves the accuracy of 65.66%:

```
params = {
    'feature_fraction': 0.65,
    'learning_rate': 0.025,
    'max_depth': 40,
    'min_data_in_leaf': 20,
    'n_estimators': 200,
    'num_leaves': 71,
}
```

It is interesting to notice that, even with many efforts to find better hyperparameters, the accuracy has not been improved too much above 65%.

## 2.2 Streak Number Test Confirms Hot Hand

We apply our streak number test on the same 153 players who took at least 1,000 shots over the two seasons with the same data as Pelechrinis and Winston use. For  $k = 4$ , with a significance level of 5%, we find that 32 out of 153 players with p-value less than 5%. Here are the best 32 players with extraordinary 4-streaks.

Table 1: Players with high number of 4-Streaks, p-value < 5%

Player	No. games	Obs 4-streaks	Exp 4-streaks	p-value
Jordan Hill	106	61	20.6	0.12%
Ersan Ilyasova	88	39	13.4	0.47%
Kemba Walker	59	32	11.1	0.79%
Tristan Thompson	140	53	22.5	0.83%
Jose Juan Barea	124	20	6.5	1.06%
Ben McLemore	124	31	10.3	1.14%
Aaron Brooks	98	26	9.0	1.22%
Kevin Durant	64	87	37.0	1.23%
Marcin Gortat	137	99	41.0	1.30%
Jeremy Lin	97	26	9.7	1.36%
Marreese Speights	117	35	11.4	1.37%
Monta Ellis	109	71	29.5	1.66%
Jonas Valanciunas	141	74	30.1	1.75%
Tony Parker	108	76	35.6	1.90%
Boris Diaw	134	43	17.2	1.93%
Deron Williams	98	32	12.3	2.09%
CJ Miles	91	26	9.3	2.10%
Kenneth Faried	127	77	36.2	2.70%
Rudy Gay	100	71	33.6	2.76%
Timofey Mozgov	143	51	22.0	2.82%
Khris Middleton	124	47	20.9	2.84%
David West	124	61	26.8	2.92%
Danny Green	111	24	8.8	3.02%
Chris Bosh	104	74	32.6	3.05%
Enes Kanter	136	80	34.5	3.37%
Markieff Morris	114	53	22.4	3.48%
Chandler Parsons	118	64	30.5	3.75%
DeMarre Carroll	114	39	17.8	3.89%
Matt Barnes	121	28	11.5	3.97%
David Lee	95	70	37.8	4.27%
Joakim Noah	121	35	15.7	4.63%
Dwight Howard	96	92	47.9	4.91%

With a significant level of 5%, the following table shows the number of players whose p-value less than the significant level for different streak numbers. For example, there are 105 out of 153 players with extraordinarily high number (p-value < 5%) of 2-streaks. With total of 153 qualified players in consideration here, there is a substantial evidence to support the existence of hot hand phenomenon in real NBA games.

Table 2: Number of Players with Hot Hand

Streak Number	# of Players with p-value < 5%	% of Total Number of Players
2	105	68.6%
3	59	38.6%
4	32	20.9%
5	22	14.4%
6	15	9.8%

### 2.3 Hot Hand Ranking

Pelechrinis and Winston develop their hot hand index based on  $P[M|k \text{ Makes}]_{Obs} - P[M|k \text{ Makes}]_{Sim}$ . A higher value means higher correlation among the player's shots, hence a hotter hand.

Our p-value associated with k-streaks can be used as another hot hand index. The smaller the p-value, the higher the player's number of streaks exceeding that of the case of independent shots, and hence the stronger display of hot hand. For both the index of Pelechrinis and Winston and ours, we calculate the weighted sum for different streak numbers weighted by the number of k-streaks. Then, we can rank the players based on the two indices. Below are the top 10 players with hot hand in both ranks.

Table 3: Top 10 Players with Hot Hand

Player	PW Rank	Our Rank	Player	Our Rank	PW Rank
Kemba Walker	1	19	Ben McLemore	1	3
Jordan Hill	2	2	Jordan Hill	2	2
Ben McLemore	3	1	Marreese Speights	3	59
Boris Diaw	4	10	Tristan Thompson	4	7
CJ Miles	5	9	Jeremy Lin	5	38
Monta Ellis	6	21	Aaron Brooks	6	36
Tristan Thompson	7	4	Markieff Morris	7	15
Chris Bosh	8	20	DeMarre Carroll	8	22
Kyle Korver	9	53	CJ Miles	9	5
Rudy Gay	10	16	Boris Diaw	10	4

Our rank of hot hand players and Pelechrinis and Winston's rank are different but mostly consistent. Their Spearman's rank correlation is 0.81.

## 2.4 ChatGPT's view of hot hand

Hot hand phenomenon is a relatively popular topic in sports. It would be interesting to see how LLMs view this issue. We explore how ChatGPT responds to the prompts about hot hand in NBA games. We use ChatGPT 5.1 Think version for the test.

Prompt:

An NBA player can have a hot hand during a game. What does having a hot hand mean?

ChatGPT:

In plain basketball language, “**having a hot hand**” means:

The player seems to “heat up” so that **after they make a few shots, they’re more likely than usual to make the next one.**

So it’s not just *shooting well overall*—it’s the idea that **makes tend to come in streaks because success feeds more success.**

More precisely, in a statistical sense, a player has a hot hand if:

- The **probability of making the next shot is higher after recent makes** than after recent misses (or higher than their usual average),
- Which shows up as **positive correlation between shots over time** (1s and 0s not being independent).

Example with 1 = make, 0 = miss:

- “Normal” but not hot: 1010010101 (about 50% makes, no long streaks, no pattern)
- Hot hand-ish: 001111011110 (same overall percentage maybe, but **lots of runs like 1111**)

So:

- **Fans’ intuition:** “He’s on fire; every shot is going in now.”
- **Statistical version:** “Given that he just made shots, the chance of making the next one is **higher** than usual or higher than after misses.”

As we can see from above answer, ChatGPT can accurately state the statistical meaning of hot hand. Next, we upload the shooting records of two players and let ChatGPT to detect which player has a hot hand. Player 1, Marreese Speights, has 117 game records and Player 2, Kyle Korver, has 128 game records.

Prompt:

Using the above criterion, i.e., a player has a hot hand if the probability of making the next shot is higher after recent makes than after recent misses (or higher than their usual average), which player do you think has a hot hand?

Below are the shooting records of two players during the same season.

Player 1: Marreese Speights

GAME\_ID,MATCHUP,SHOT\_NUMBER  
21300225 , NOV 27 2013 - GSW @ DAL , 1  
21300738 , FEB 06 2014 - GSW vs. CHI , 0101  
21300355 , DEC 15 2013 - GSW @ PHX , 01010  
...  
21400922 , MAR 06 2015 - GSW vs. DAL , 11000110  
21400426 , DEC 23 2014 - GSW @ LAL , 0110010  
21400703 , JAN 30 2015 - GSW @ UTA , 000001

Player 2: Kyle Korver

GAME\_ID,MATCHUP,SHOT\_NUMBER  
21301004 , MAR 18 2014 - ATL vs. TOR , 00001111001000  
21301120 , APR 02 2014 - ATL vs. CHI , 110100001  
21300483 , JAN 03 2014 - ATL vs. GSW , 1001000  
...  
21401138 , APR 04 2015 - ATL vs. BKN , 101001  
21400958 , MAR 11 2015 - ATL @ DEN , 10011111  
21400319 , DEC 10 2014 - ATL vs. PHI , 01110011010

ChatGPT:

Using your criterion (“hot hand” = higher chance of making the next shot after a make than after a miss / than overall), Player 1 **Marreese Speights** looks more like he has a hot hand than Kyle Korver.

Here’s what happens if we look at *within-game* consecutive shots (i.e., only transitions from one shot to the very next shot in the same game):

### Marreese Speights (Player 1)

- Total shots: 766
- Makes: 361  $\Rightarrow$  overall make %  $\approx 47.1\%$

Conditioning on the previous shot:

- $\mathbb{P}(\text{make} \mid \text{previous} = \text{make}) \approx 48.1\%$
- $\mathbb{P}(\text{make} \mid \text{previous} = \text{miss}) \approx 45.8\%$

So his chance of making the next shot is **higher after a make than after a miss**, and slightly higher than his overall average — this matches your “hot hand” criterion (at least directionally).

### Kyle Korver (Player 2)

- Total shots: 1015
- Makes: 493  $\Rightarrow$  overall make %  $\approx 48.6\%$

Conditioning on the previous shot:

- $\mathbb{P}(\text{make} \mid \text{previous} = \text{make}) \approx 48.4\%$
- $\mathbb{P}(\text{make} \mid \text{previous} = \text{miss}) \approx 48.1\%$

Those numbers are **almost identical** and very close to his overall average, which is what you’d expect if his shots were basically independent.

## Conclusion

- By your definition, **Speights** shows a small “hot-hand-like” effect (better after makes than after misses).
- **Korver** does not: his make probability barely changes based on what happened on the previous shot.

As we can see from the answer, ChatGPT can perform the calculation of conditional probability  $P[M \mid 1 \text{ Make}]$  and  $P[M \mid 1 \text{ Miss}]$  and uses them to detect hot hand. We test ChatGPT on total 4 pairs of players.

- Ben McLemore (our rank = 1, PW rank = 3) vs Kemba Walker (our rank = 19, PW rank = 1). Our Number 1 player vs PW Number 1 player.  
ChatGPT: Ben McLemore has hot hand
- Marreese Speights (our rank = 3, PW rank = 59) vs Kyle Korver (our rank = 53, PW rank = 9). Two inconsistently ranked players.  
ChatGPT: Marreese Speights has hot hand
- CJ Miles (our rank = 9, PW rank = 5) vs Ersan Ilyasova (our rank = 24, PW rank = 26). Two consistently ranked players.  
ChatGPT: CJ Miles has hot hand
- Jeremy Lin (our rank = 5, PW rank = 38) vs Chris Bosh (our rank = 20, PW rank = 8). Two moderately inconsistently ranked players.  
ChatGPT: Jeremy Lin has hot hand

Overall, ChatGPT tends to align more with our rank than Pelechrinis and Winston’s rank. The player who ranks higher in our rankings is viewed by ChatGPT as having a hotter hand.

### 3. Conclusion

The hot hand phenomenon is a very interesting and important topic in sports and other fields. It concerns whether there is a positive correlation among the outcomes of repeated human performance under stress. Advances in data science provide more powerful tools for our investigations. Through the project,

1. We develop a shot-make prediction model using LightGBM that achieves a better accuracy than Pelechrinis and Winston's deep learning model.
2. We develop a new statistical test - streak number test - and find strong evidence that the hot hand exists in NBA games.
3. We test ChatGPT's responses to hot hand questions and find it can correctly identify the statistical meaning of hot hand and perform some statistical estimation on real data.

### 4. Workflow

We all contribute to the project. Peter forms the idea of the hot hand project, he and Leon work more on the streak selection bias and shot-make models. Yuan and Leon work more on the streak number test and ChatGPT's responses to hot hand prompts.

1. Week of 10/12: Study the papers (Peter, Leon, Yuan)
2. Week of 10/19: Study the papers, examine the data, and results (Peter, Leon, Yuan)
3. Week of 10/26: Build models, prepare the data (Peter, Leon, Yuan)
4. Week of 11/02: Build shot-make prediction model (Peter, Leon)
5. Week of 11/02: Test models, analyze the results (Peter, Leon, Yuan)
6. Week of 11/02: Organize the results, prepare Milestone 2 (Peter, Leon, Yuan)
7. Week of 11/09: Build streak number test models (Leon, Yuan)
8. Week of 11/16: Study if we can obtain additional data (Peter, Yuan)
9. Week of 11/23: Test ChatGPT's responses (Leon, Yuan)
10. Week of 11/30: Summarize works and prepare slides (Peter, Leon, Yuan)
11. Week of 12/07: Wrap up, write Milestone 3 (Peter, Leon, Yuan)

### References

Vallone R. Gilovich, T. and A. Tversky. The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3):295–314, 1985.

J. B. Miller and A. Sanjurjo. Surprised by the hot hand fallacy? a truth in the law of small numbers. *Econometrica*, 86(6):2019–2047, 2018.

K. Pelechrinis and W. Winston. The hot hand in the wild. *PLoS ONE*, 17(1):e0261890, 2022.