

7106056041 靳開翔  
7106056033 張家豪  
7106056109 江岳勳

Social Network Final Project

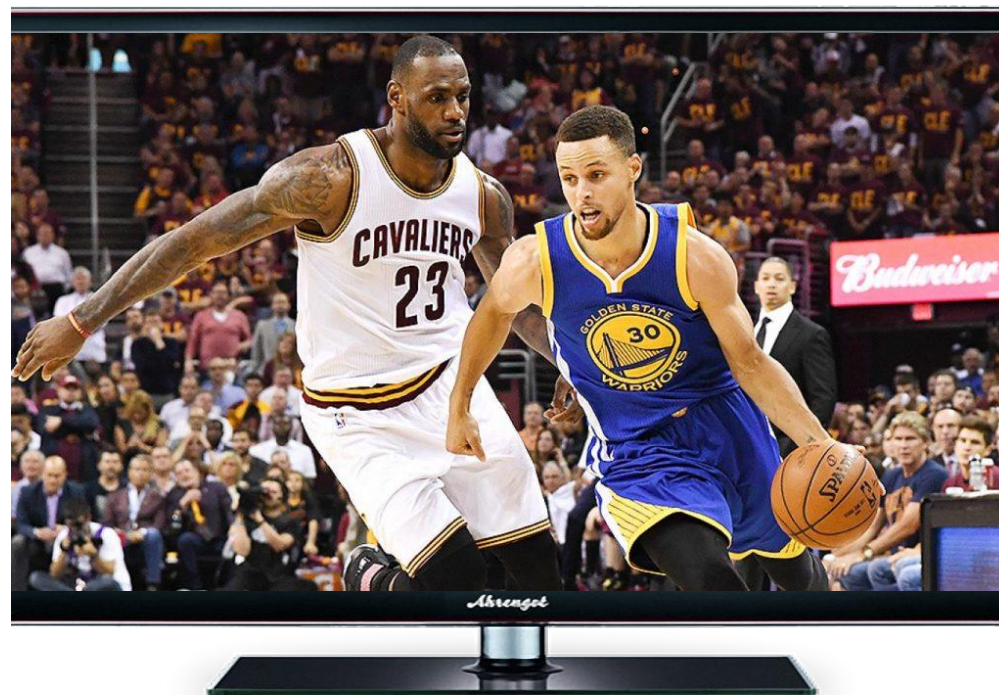
# NBA鄉民來預測





NBA球賽進行時，網友會到討論區談論比賽狀況  
網友們的討論是否與實際狀況符合？  
能否從中發掘到特別的關係、現象？

## 研究動機







# 研究目標(1/2)

資料集: PTT 論壇中 NBA 比賽的直播文  
球員表現的討論度量化為預測勝負與比分的指標

推 stanmax9999 : Curry for 333333333333333333333333  
推 Strasburg : KK怎麼穿得像流浪漢  
→ rayes : 嘻嘻 又在總亞軍 只能活在過去真慘  
推 vestal : 騎士需要有人跳出來戳kd的眼睛  
推 gn1384181 : 今天這場可能是母獅最後一次穿騎士的球衣  
推 i945566 : LBJ眼睛還是紅的  
推 a5566691 : 開啟血輪眼了  
推 a6964528 : 姆斯眼睛有夠紅  
推 ppo7741 : 前面兩個球評根本沒在回答正磊的問題阿笑死都在講  
→ ppo7741 : 自己的  
推 whyhsu : KK帥  
推 King5566 : 唉 都不覺得騎士會贏了  
推 qazzaq3977 : 軍人也JR  
推 applekiwi : 嘴綠插眼關係?  
→ hunt5566 : 抓到囉 慕斯眼紅勇士冠軍  
推 SCEW : 演唱嘉賓預測FMVP了



Curry +1



## 研究目標(2/2)

從球員表現預測

1. 球隊總得分
2. 勝負結果
3. 賭盤



## 前處理(1/4)

爬蟲抓取 NBA 版中特定球賽的直播文  
EX: 季後系列賽程

勇士 VS 火箭 GAME1

.

.

.

.

.

.

勇士 VS 火箭 GAME7



## 前處理(2/4)

別名問題:

一個NBA球星可能有多種稱呼、綽號

EX:

Stephen Curry:

Curry

咖哩

路人

科呂

浪花兄...



## 前處理(3/4)

當鄉民討論到某個球員時不一定是他表現好:

→ jackman6666 : green今天在幹嘛  
推 Costco5566 : 看推文以為勇4被血虐  
推 Klaver : 第二年分區冠就打這麼抖 歷史第一強隊?  
推 tallolz : 超快節奏  
推 nedsonyy : green傳球真的爛到笑

當失誤或犯規行為出現時  
對球員的評價視為反向指標



## 前處理(4/4)

轉換完成表格(含得分與勝負結果)

game	team	coach	F	F	C	G	G	o1	o2	o3	o4	o5	o6	o7	o8	o9	score	result
0	112	11	60	93	0	17	42	15	13	15	5	0	1	58	6	4	113	1
1	113	4	78	61	1	27	87	60	5	8	3	0	3	8	8	15	116	1
2	80	1	33	14	2	13	26	6	10	4	13	0	2	4	2	13	110	1
3	79	0	28	12	0	7	15	4	1	0	1	0	4	1	0	1	90	0
4	150	11	146	48	1	48	77	3	16	2	11	0	0	3	1	5	99	1
5	178	8	98	30	1	51	56	8	10	32	8	0	0	6	0	13	123	1
6	150	10	86	31	0	200	133	8	0	2	16	0	0	2	0	1	121	1
7	207	6	34	49	0	54	80	4	0	3	1	0	0	4	0	61	100	0
8	88	8	84	38	0	45	17	22	8	0	4	0	2	3	3	10	118	1
9	142	0	54	55	0	62	70	7	9	0	2	0	0	0	0	31	113	1
10	239	1	99	91	0	76	18	7	25	9	0	0	0	0	0	20	119	1
11	228	6	104	35	0	73	22	4	7	6	12	0	0	0	0	67	105	0
12	196	2	62	15	5	154	10	3	9	4	2	0	1	2	3	0	126	1
13	158	20	45	39	7	68	25	13	4	3	14	0	0	0	0	1	92	0
14	295	10	154	60	1	87	34	11	2	0	0	0	0	2	1	1	94	0
15	276	10	138	25	4	44	71	8	6	10	0	2	0	5	0	17	115	1
16	172	3	87	32	3	33	90	3	1	8	2	0	0	2	0	2	101	1





## Datasets Info

- 隊伍名稱
- 教練
- 先發五名球員
- 替補九名成員
- 基本欄位共16個，判對好壞後共32個特徵
- 預測得分
- 獲勝隊伍
- 賭盤比分



# 機器學習方法

Logistic regression

Linear regression

Random forest



**GAME1** Warriors Win

**GAME2** Warriors Win

**GAME3** Warriors Win

**GAME4** Warriors Win

2018 NBA finals



<b>GAME1</b>	<b>Warriors WIN</b>
<b>GAME2</b>	<b>Rockets WIN</b>
<b>GAME3</b>	<b>Warriors WIN</b>
<b>GAME4</b>	<b>Rockets WIN</b>
<b>GAME5</b>	<b>Rockets WIN</b>
<b>GAME6</b>	<b>Warriors WIN</b>
<b>GAME7</b>	<b>Warriors WIN</b>

2018 Western Conference Finals

## 最終預測(1/3)

	火箭得分	火箭預測	勇士得分	勇士預測	實際勝負	預測勝負
GAME 5	98	95	94	92	火箭WIN	火箭WIN
GAME 6	86	85	115	126	勇士WIN	勇士WIN
GAME 7	92	95	101	92	勇士WIN	勇士WIN

# 最終預測(2/3)

總冠軍戰	隊伍	比數	結果	分差	總得分	運彩讓分	運彩大小分	國際讓分	國際大小分
Game 1	勇士	124	Win	10	238	-12.5	217.5 大	-13	217.5 大
	騎士	114							
Game 2	勇士	122	Win	19	225	-11.5	215.5 大	-11.5	216.5 大
	騎士	103							
Game 3	勇士	110	Win	8	212	-3.5	216.5 小	-4.5	216 大
	騎士	102							
Game 4	勇士	108	Win	23	193	-3.5	212.5 小	-4.5	216.5 小
	騎士	85							
Game 4 pred	勇士	122	Win	19	225		小(N)		
	騎士	103							



# 最終預測(3/3)

西部決賽	隊伍	比數	結果	分差	總得分	運彩讓分	運彩大小分	國際讓分	國際大小分
Game 1	勇士	119	Win	13	225	2.5	208.5 小	1.5	225.5 小
	火箭	106							
Game 2	勇士	105		-22	232	1.5	214.5 大	2	225 大
	火箭	127	Win						
Game 3	勇士	126	Win	41	211	-7.5	220.5 小	-7.5	226.5 小
	火箭	85							
Game 4	勇士	92		-3	187	-8.5	226.5 小	-8	227 小
	火箭	95	Win						
Game 5	勇士	94		-4	192	-1.5	226.5 小	1.5	220 大
	火箭	98	Win						
Game 6	勇士	115	Win	29	201	-12.5	224.5 小	-12.5	214.5 小
	火箭	86							
Game 7	勇士	101	Win	9	193	-6.5	225.5 小	-6	208.5 小
	火箭	92							
Game 5 pred	勇士	92		-3	187	N	小(Y)		
	火箭	95	Win						
Game 6 pred	勇士	126	Win	41	211	Y	小(Y)		
	火箭	85							
Game 7 pred	勇士	92							
	火箭	95	Win	-3	187	N	小(Y)		

## 困難的部分

- 深夜時段比賽沒人討論
- 不同的對決收視觀眾不同
- 一篇文章中有多少個推文
- 多少個推文在有提到人名

## Future Work

- 與傳統方式比較
  - 傳統作法以實際球員場上表現為準
- 推文時間區段限制
  - 短時間內有太多人對同一人評價



# 參考文獻

- [Social Media Data Mining on 2017 NBA Champ](#)
- [Data Mining Final ProjectRedditTextMining](#)