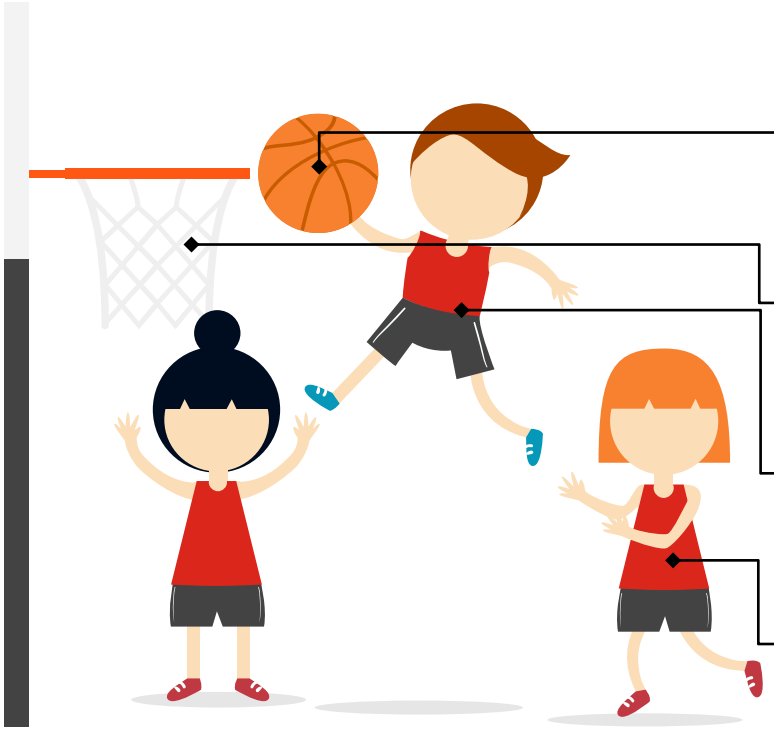
Dunking the Winning Formula

Predicting NBA playoff contention from team stats

NBA Team Stats Season 2021-2022

- Taken from Basketball Reference
- Manually compiled with love and care by yours truly
- Includes 30 rows (1 per team) and 61 columns
- Has four feature categories
- Keyword being “season” so it only involves regular season performance

Scope and Limitations



1. The game is ever-changing and evolving, officiating and trending playstyles are different per era of basketball.
2. If major rule changes are to take place right before a season that is to be subject to prediction given these methods, insights may lose weight.
3. This kind of method of prediction may move differently in seasons where analytics were less evolved, with less of these stats introduced then.
4. In the same way, should new derived stats be introduced, results could change as well.

Feature Categories

Basic Stats	Advanced Stats	Shooting Stats	Predictive Stats
Points per game	Average age	Average shot distance	Predicted wins
Field goal accuracy	Offensive and Defensive Rating	Percentage of field goals attempted from 0 to 3, 3-10 and 10-16 feet	Predicted losses
Assists per game	Average possessions per 48 minutes	Dunks made	
Total rebounds per game	True shooting percentage	Layups made	

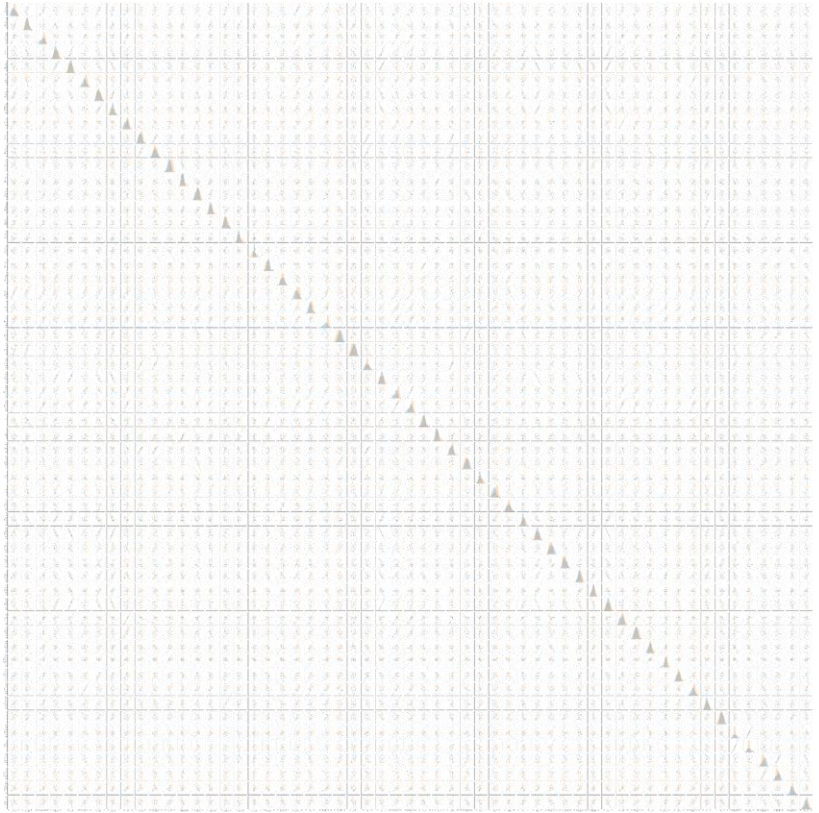
Features

Variable	Description
Team	Self-explanatory
Made playoffs?	Whether a team made the playoffs or not
FG per game	Field goals made per game
FGA per game	Field goals attempted per game
3P per game	3-pointers made per game
3PA per game	3-pointer attempted per game
3P%	3-pointer accuracy
2P per game	2-pointers made per game
2PA per game	2-pointer attempted per game
2P%	2-pointer accuracy
FT per game	Free throws made per game
FTA per game	Free throws attempted per game
FT%	Free throw accuracy
ORB	Offensive rebounds per game
DRB	Defensive rebounds per game
TRB	Total rebounds per game
AST per game	Assists per game
STL per game	Steals per game
BLK per game	Blocks per game
TOV per game	Turnovers per game
PF per game	Personal fouls per game
PTS per game	Points per game
Age	Average age of players in a team
W	Wins
L	Losses
PW	Predicted Wins
PL	Predicted Losses
MOV	Average Margin of Victory
SOS	Strength of Schedule
SRS	Simple Rating System (Average point differential and strength of schedule)
ORtg	Offensive Rating (Estimated points produced per 100 possessions)
DRtg	Defensive Rating (Estimated points allowed per 100 possessions)
NRtg	Net Rating (Estimate point differential per 100 possessions)
Pace	Estimate possessions per 48 minutes
FTr	Free Throw Attempt Rate
3PAr	Three Point Attempt Rate

AST per game	Assists per game
TS%	Shooting efficiency accounting for 2-pointers, 3-pointers and free throws
O-eFG%	Field goal percentage except 3-pointers are given more weight since they're worth more points made by team
O-TOV%	Turnover Percentage made by team
ORB%	Offensive Rebound Percentage
O-FT/FGA	Free throws per field goal attempt by team
D-eFG%	Field goal percentage except 3-pointers are given more weight since they're worth more points allowed by team
D-TOV%	Turnover Percentage allowed by team
DRB%	Offensive Rebound Percentage allowed by team
D-FT/FGA	Free throws per field goal attempt allowed by team
Dist.	Average distance per field goal attempt
2P FGA%	Percent of field goals attempted that are 2-pointers
0-3 ft FGA%	Percent of field goals attempted attempted 0-3 feet away from the basket
3-10 ft FGA%	Percent of field goals attempted attempted 3-10 feet away from the basket
10-16 ft FGA%	Percent of field goals attempted attempted 10-16 feet away from the basket
16-3P ft FGA%	Percent of field goals attempted attempted 16 feet to 3-point range from the basket
3P FGA%	Percent of field goals attempted from 3-point range
2P FG%	Percent of field goals attempted that are 2-pointers
0-3 ft FG%	Accuracy of field goals attempted attempted 0-3 feet away from the basket
3-10 ft FG%	Accuracy of field goals attempted attempted 3-10 feet away from the basket
10-16 ft FG%	Accuracy of field goals attempted attempted 10-16 feet away from the basket
16-3P ft FG%	Accuracy of field goals attempted attempted 16 feet to 3-point range from the basket
3P FG%	Accuracy of field goals attempted from 3-point range
2P	Percentage of 2-pointers that were assisted
3P	Percentage of 3-pointers that were assisted
Dunks %FGA	Percent of field goals attempted that are dunks
Dunks Md.	Total number of dunks made
Layups %FGA	Percent of field goals attempted that are layups
Layups Md.	Total number of layups made
Corner %3PA	Percent of field goals attempted that are corner three-pointers
Corner %3P	Accuracy of corner three-pointer attempts

Collinearity Analysis

Same, LeBron. Same.



More obvious collinearities

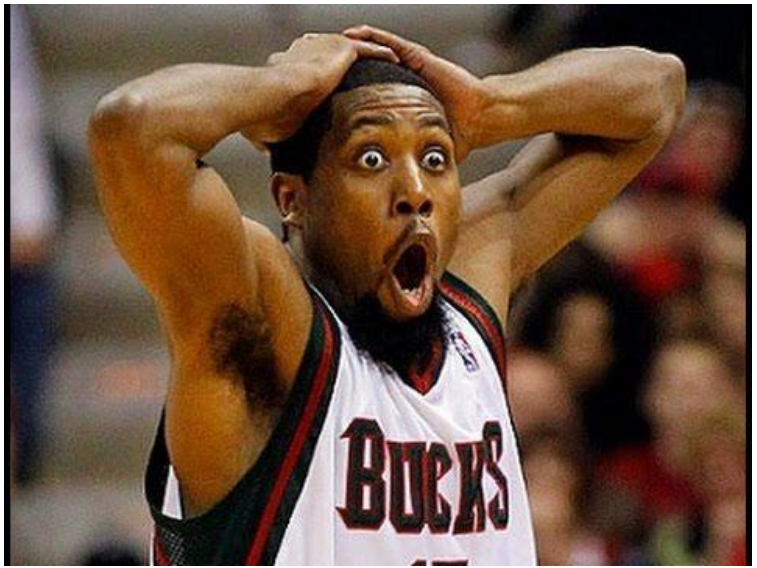
- The more two-pointers you attempt, the higher your 2P FGA% and 2-pointers made are; same applies with three-pointers and related stats
- The higher your 3P FGA%, the lower your 2P FGA% is and 2-pointers attempted and made are
- The more attempts you have, the more makes you'll have (That's how volume shooting works)
- General accuracy stats are collinear with derived accuracy stats such as TS% and eFG%
- Higher average distance means less 0-3 and 3-10 ft FGA% and more 10-16 and 16-3P FGA%, it can also mean more three-pointers attempted per game

Intriguing collinearities and the lack of thereof

- Age average and FG%: May suggest with age comes both maturity and wiser shot selection and therefore a more refined skillset

- Turnovers and steals not having collinearity: Turnovers don't always translate to steals, this is because it could happen from committing violations and stepping out of bounds

- Steals and dunks/layups not having collinearity: Steals don't always translate to dunks or layups, converting points out of turnovers is an art and entirely new challenge in itself - transitions in between playing offense and defense are important



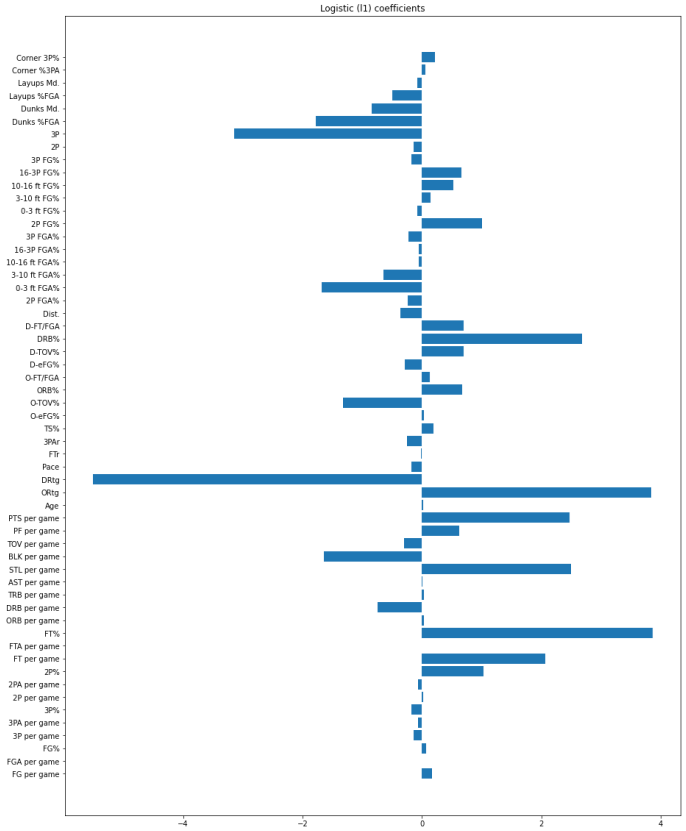
Model Tests

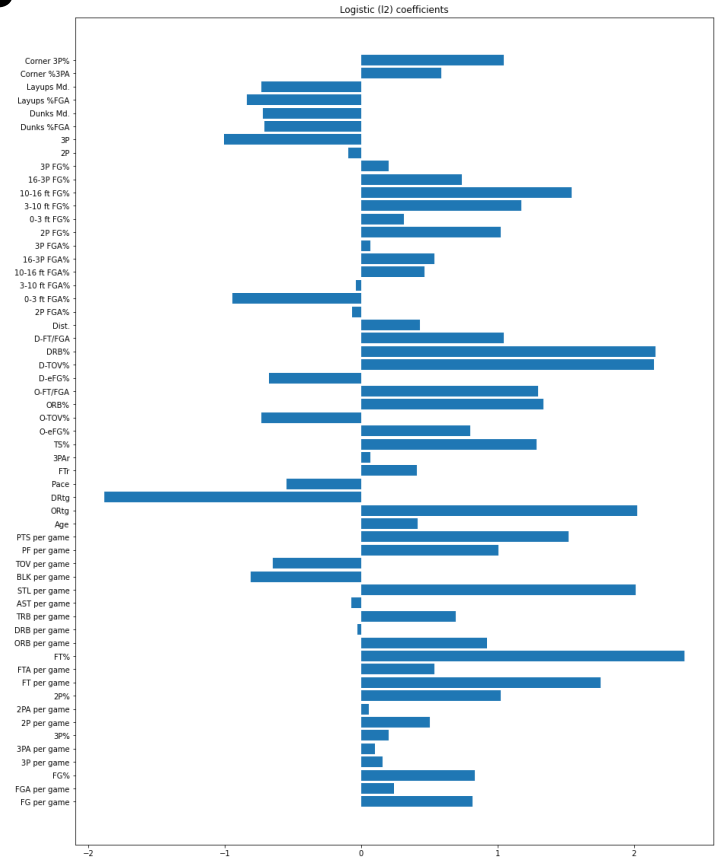
All Stats

- DRtg and ORtg reign supreme over other predictor variables with FT% edging the others in one of the models. DRtg and ORtg measure defensive and offensive potency, so having better values for both would lead to more success

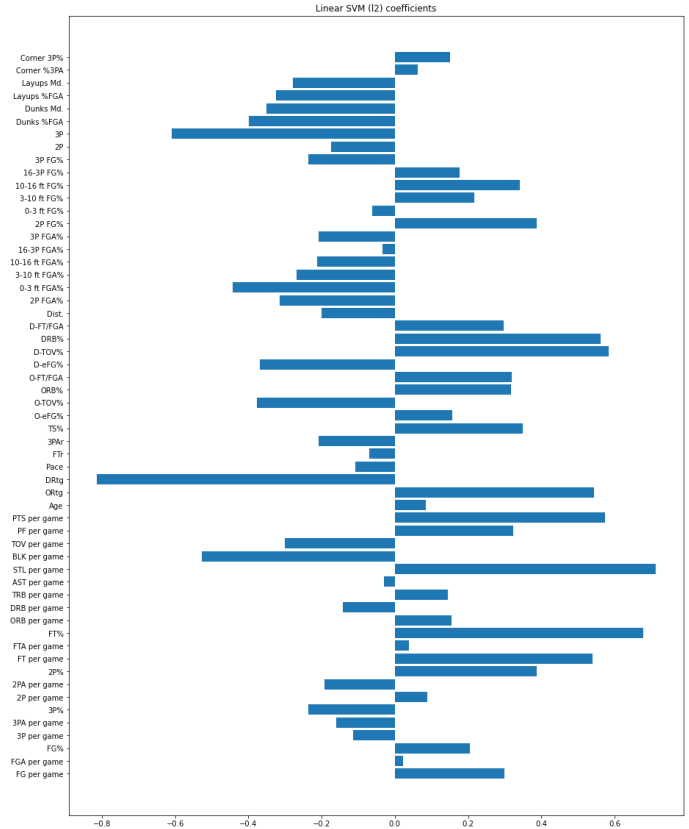
Machine Learning Method	Test Accuracy	Best Parameter	Top Predictor Variable	Run Time
kNN	0.741667	N_Neighbor = 4	NA	132.707502
Logistic (l2)	0.854167	C = 300	FT%	3.395071
Logistic (l1)	0.783333	C = 5000	DRtg	1.703554
LinearSVM (l2)	0.891667	C = 300	DRtg	1.830221
LinearSVM (l1)	0.845833	C = 5000	DRtg	5.322011
Random Forest	0.770833	Max_features = 0.4	ORtg	32.104800
Gradient Boosting Method	0.770833	Max_features = 0.5	ORtg	11.186385

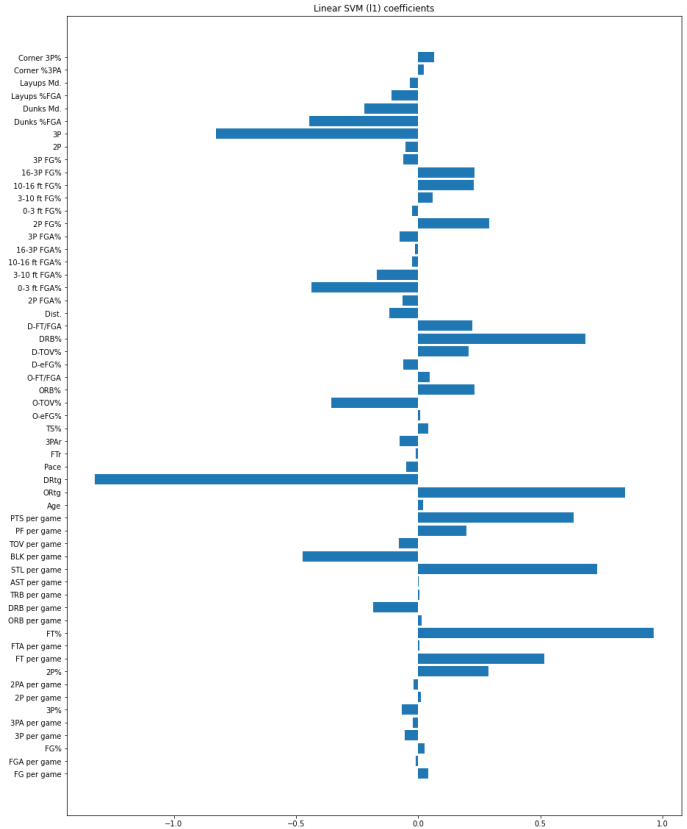
All Stats


Logistic (l1) coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
−4
−2
0
2
4

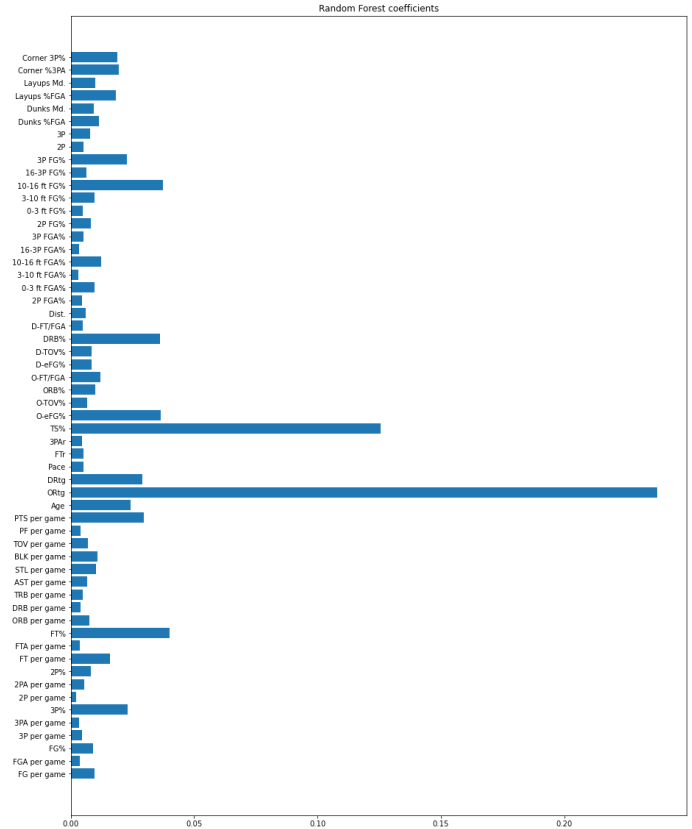

Logistic (l2) coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
−2
−1
0
1
2

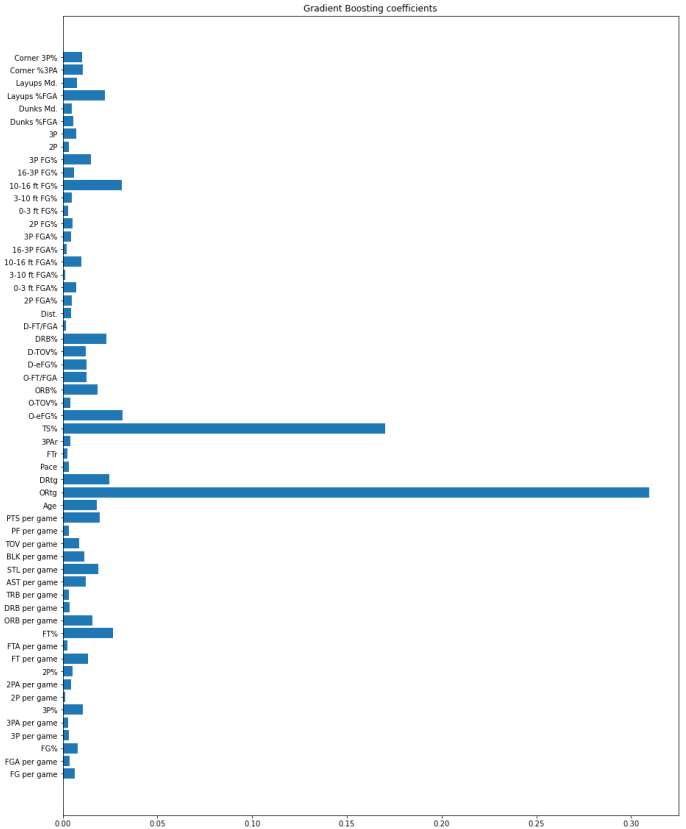
All Stats


Linear SVM (l2) coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6


Linear SVM (l1) coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
−1.0
−0.5
0.0
0.5
1.0

All Stats


Random Forest coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
0.00
0.05
0.10
0.15
0.20


Gradient Boosting coefficients
Corner 3P%
Corner %3PA
Layups Md.
Layups %FGA
Dunks Md.
Dunks %FGA
3P
2P
3P FG%
16-3P FG%
10-16 ft FG%
3-10 ft FG%
0-3 ft FG%
2P FG%
3P FGA%
16-3P FGA%
10-16 ft FGA%
3-10 ft FGA%
0-3 ft FGA%
2P FGA%
Dist.
D-FT/FGA
DRB%
D-TOV%
D-eFG%
O-FT/FGA
ORB%
O-TOV%
O-eFG%
TS%
3PAr
FTr
Pace
DRtg
ORtg
Age
PTS per game
PF per game
TOV per game
BLK per game
STL per game
AST per game
TRB per game
DRB per game
ORB per game
FT%
FTA per game
FT per game
2P%
2PA per game
2P per game
3P%
3PA per game
3P per game
FG%
FGA per game
FG per game
0.00
0.05
0.10
0.15
0.20
0.25
0.30

All stats without ORtg, DRtg and PPG

Machine Learning Method	Test Accuracy	Best Parameter	Top Predictor Variable	Run Time
kNN	0.733333	N_Neighbor = 2	NA	134.230604
Logistic (l2)	0.837500	C = 5000	FT%	3.289656
Logistic (l1)	0.783333	C = 5000	FT%	1.647990
LinearSVM (l2)	0.850000	C = 10	STL per game	1.838235
LinearSVM (l1)	0.825000	C = 5000	FT%	5.562322
Random Forest	0.687500	Max_features = 0.7	TS%	31.935145
Gradient Boosting Method	0.741667	Max_features = 0.8	TS%	11.603374

Shooting Stats

Machine Learning Method	Test Accuracy	Best Parameter	Top Predictor Variable	Run Time
kNN	0.662500	N_Neighbor = 2	NA	139.215353
Logistic (l2)	0.637500	C = 1	10-16 ft FG%	3.917897
Logistic (l1)	0.570833	C = 3	3P	1.770935
LinearSVM (l2)	0.554167	C = 1	10-16 ft FG%	1.907589
LinearSVM (l1)	0.562500	C = 1	10-16 ft FG%	7.074812
Random Forest	0.683333	Max_features = 0.1	3P FG%	33.306932
Gradient Boosting Method	0.666667	Max_features = 0.1	3P FG%	12.366834

Basic Stats

Machine Learning Method	Test Accuracy	Best Parameter	Top Predictor Variable	Run Time
kNN	0.716667	N_Neighbor = 5	NA	119.630021
Logistic (l2)	0.825000	C = 300	FT%	2.737521
Logistic (l1)	0.775000	C = 100	FT%	1.434030
LinearSVM (l2)	0.833333	C = 3	FT%	1.525958
LinearSVM (l1)	0.800000	C = 5000	FT%	4.447250
Random Forest	0.616667	Max_features = 0.7	FT%	31.185139
Gradient Boosting Method	0.620833	Max_features = 0.2	FT%	11.496053

Advanced Stats

Age ORtg DRtg Pace FTr 3PAr TS% O-eFG% O-TOV% ORB% O-FT/FGA D-eFG% D-TOV% DRB% D-FT/FGA

- Offensive rating taking precedence when only advanced stats are included in the model testing but defensive rating narrowly takes precedence when all stats are accounted for.

Machine Learning Method	Test Accuracy	Best Parameter	Top Predictor Variable	Run Time
kNN	0.775000	N_Neighbor = 11	NA	68.791311
Logistic (l2)	0.812500	C = 1000	ORtg	2.645219
Logistic (l1)	0.775000	C = 15	ORtg	1.398915
LinearSVM (l2)	0.812500	C = 100	ORtg	1.387224
LinearSVM (l1)	0.775000	C = 5	ORtg	3.250083
Random Forest	0.854167	Max_features = 0.7	ORtg	30.893874
Gradient Boosting Method	0.820833	Max_features = 0.3	ORtg	10.441808

Advanced Stats

- Even going by eye-tests, you can see the clear correlation between regular season success and offensive/defensive rating*

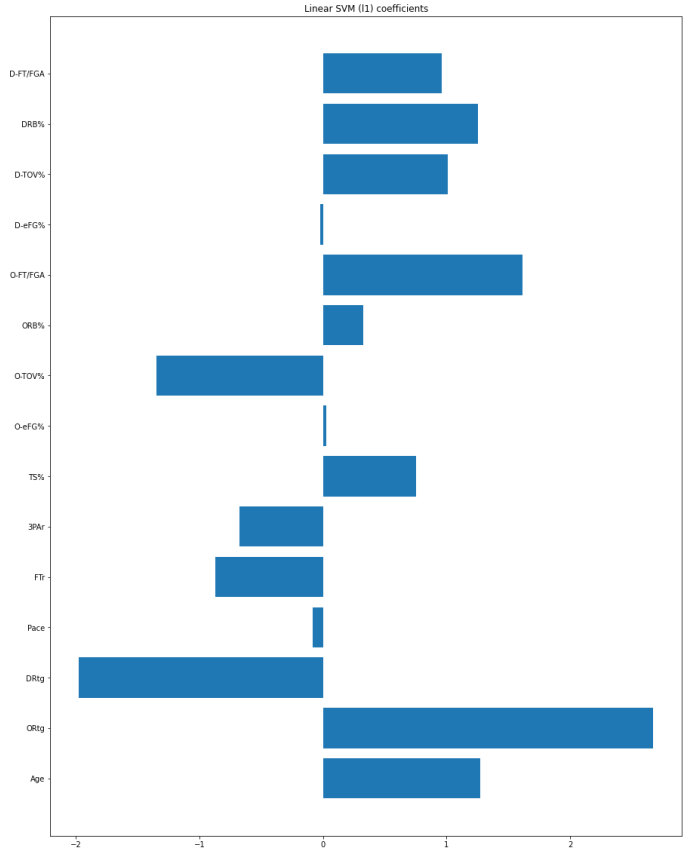
*Note: Offensive rating is better the higher it is, defensive rating better the lower it is

Rk	Team	Age	W	L	PW	PL	MOV	SOS	SRS	ORtg	DRtg	NRtg	Pace	FTr	3PAr	TS%
1	Boston Celtics*	26.1	51	31	59	23	7.28	-0.26	7.02	114.4	106.9	+7.5	96.6	.239	.425	.578
2	Phoenix Suns*	27.5	64	18	59	23	7.50	-0.56	6.94	114.8	107.3	+7.5	99.8	.221	.354	.581
3	Utah Jazz*	29.3	49	33	56	26	6.04	-0.37	5.67	116.7	110.5	+6.2	97.1	.271	.468	.589
4	Memphis Grizzlies*	24.0	56	26	55	27	5.68	-0.32	5.37	114.6	109.0	+5.6	100.3	.245	.346	.553
5	Golden State Warriors*	27.6	53	29	55	27	5.54	-0.02	5.52	112.5	106.9	+5.6	98.4	.235	.456	.582
6	Miami Heat*	28.2	53	29	53	29	4.45	-0.22	4.23	113.7	109.1	+4.6	95.9	.252	.422	.584
7	Dallas Mavericks*	26.7	52	30	50	32	3.30	-0.18	3.12	112.8	109.4	+3.4	95.4	.249	.439	.572
8	Milwaukee Bucks*	28.5	51	31	49	33	3.35	-0.14	3.22	115.1	111.8	+3.3	99.9	.257	.430	.580
9	Philadelphia 76ers*	26.8	51	31	48	34	2.61	-0.04	2.57	113.5	110.8	+2.7	96.2	.282	.376	.578
10	Minnesota Timberwolves*	24.2	46	36	48	34	2.63	-0.10	2.53	114.3	111.7	+2.6	100.9	.254	.454	.573
11	Denver Nuggets*	27.7	48	34	47	35	2.30	-0.15	2.16	114.5	112.1	+2.4	97.8	.244	.416	.590
12	Toronto Raptors*	24.8	48	34	47	35	2.29	0.08	2.38	112.9	110.5	+2.4	96.0	.234	.375	.543
13	Cleveland Cavaliers	24.7	44	38	47	35	2.12	-0.08	2.04	111.9	109.7	+2.2	96.1	.261	.387	.571
14	Atlanta Hawks*	26.1	43	39	45	37	1.56	-0.01	1.55	116.5	114.9	+1.6	97.7	.253	.390	.581
15	Brooklyn Nets*	29.1	44	38	43	39	0.78	0.04	0.82	113.6	112.8	+0.8	99.0	.246	.359	.576
16	Charlotte Hornets	25.5	43	39	42	40	0.44	0.09	0.53	114.1	113.7	+0.4	100.0	.234	.418	.572
17	San Antonio Spurs	24.5	34	48	41	41	0.12	-0.10	0.02	112.4	112.3	+0.1	100.0	.220	.345	.556
18	Los Angeles Clippers	27.5	42	40	41	41	0.02	0.06	0.09	110.1	110.1	0.0	98.0	.224	.391	.564
19	New York Knicks	26.0	37	45	41	41	-0.12	0.11	-0.01	110.4	110.5	-0.1	95.9	.280	.428	.550
20	Chicago Bulls*	26.3	46	36	40	42	-0.39	0.02	-0.38	113.2	113.6	-0.4	98.3	.248	.332	.579
21	New Orleans Pelicans*	25.6	36	46	38	44	-0.99	0.15	-0.84	112.0	113.0	-1.0	97.2	.264	.365	.557
22	Los Angeles Lakers	30.2	33	49	33	49	-3.05	-0.03	-3.08	110.3	113.3	-3.0	100.1	.259	.388	.567
23	Washington Wizards	25.9	35	47	32	50	-3.38	0.15	-3.23	111.1	114.5	-3.4	97.0	.252	.356	.568
24	Indiana Pacers	26.0	25	57	32	50	-3.48	0.22	-3.26	112.6	116.1	-3.5	98.0	.239	.395	.564
25	Sacramento Kings	25.8	30	52	28	54	-5.46	0.20	-5.26	109.9	115.3	-5.4	99.8	.264	.377	.561
26	Detroit Pistons	23.6	23	59	22	60	-7.72	0.37	-7.36	106.0	113.8	-7.8	98.4	.249	.391	.533
27	Orlando Magic	23.3	22	60	21	61	-8.00	0.33	-7.67	104.5	112.5	-8.0	99.2	.223	.417	.538
28	Oklahoma City Thunder	22.4	24	58	21	61	-8.10	0.20	-7.90	104.6	112.8	-8.2	98.5	.223	.419	.530
29	Houston Rockets	24.1	20	62	21	61	-8.48	0.22	-8.26	108.4	116.7	-8.3	100.9	.284	.448	.565
30	Portland Trail Blazers	25.6	27	55	20	62	-8.88	0.33	-8.55	107.8	116.9	-9.1	98.3	.248	.422	.550
	League Average	26.3			41	41	0.00	0.00	0.00	112.0	112.0		98.2	.248	.399	.566

Advanced Stats

- Offensive and defensive rating have the most weight among all predictor variables in majority of the model tests

*Note: Offensive rating is better the higher it is, defensive rating better the lower it is



Recommendations

Refine 3P shooting

Distance shooting dominates in today's stretched NBA court

Improve defense

Defense wins games. Bill Russell won eleven rings doing so.

Practice free throws

Free throws are the easiest points in the game, make the most out of them

Balance team age

Every team needs veterans to lead and help nurture young talent