NBA Game Outcome Classifier

Reed Whitcraft
05.14.2019

The Question

Can NBA game outcomes (W/L) be accurately predicted using game data from the first three quarters?

And, if so, how accurately?

The Data

Source: stats.nba.com

Seasons: 7 Seasons (2012 - 2019)

Games: 8,000+

Features: PTS, REBS, AST, STL, BLK, TOV, FG%, +/-

Four Factors - Score, Protect, Crash & Attack

Effective Field Goal %

eFG% =

$$\frac{\text{FGM} + 0.5 * \text{3PM}}{\text{FGA}}$$

Best relative measurement for points per field goal attempt.

Offensive Rebound %

OREB% =

$$\frac{\text{OREB}}{\text{OREB} + \text{OPP DREB}}$$

Extends possession, allows for second attempt

Free Throw Rate

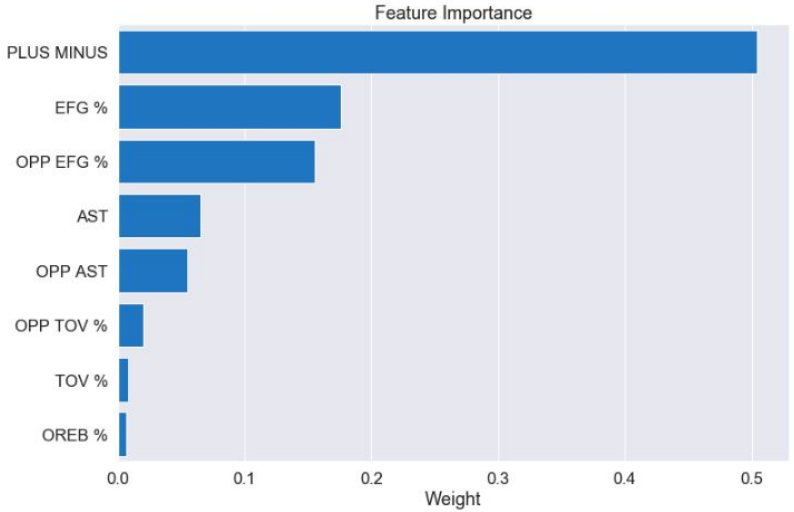
FT RATE =

$$\frac{\text{Free Throw Attempts}}{\text{Field Goal Attempts}}$$

Guaranteed attempts to score, limits opposing team's players

Feature Importance

What happened to Free Throw Rate?



Most Important: Plus/Minus

Secondary: EFG%, OPP EFG%

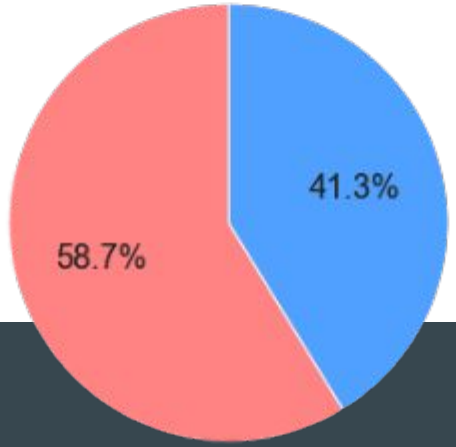
Low: AST, OPP AST

Minimal: OPP TOV%, TOV%, OREB%

WIN

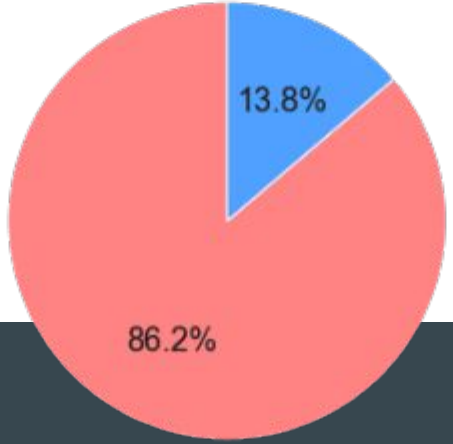
PLUS/MINUS @ 3RD QUARTER

LOSS



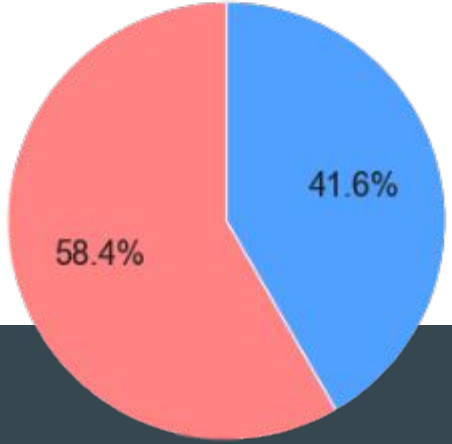
Overall W/L

Home Team won **58.7%** of the time over the last 7 seasons



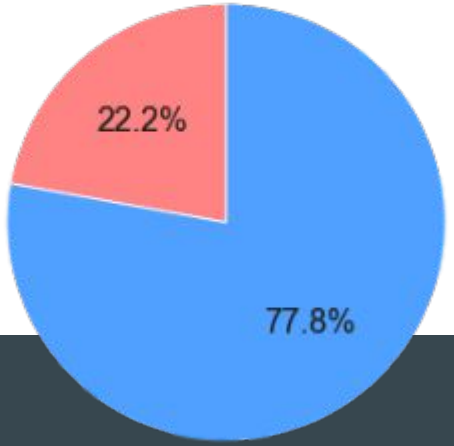
W/L (Up)

Home Team won 86.2% of the time when leading at the end of the third.



W/L (Tie)

Home Team won 58.4% of the time when tied at the end of the third.



W/L (Down)

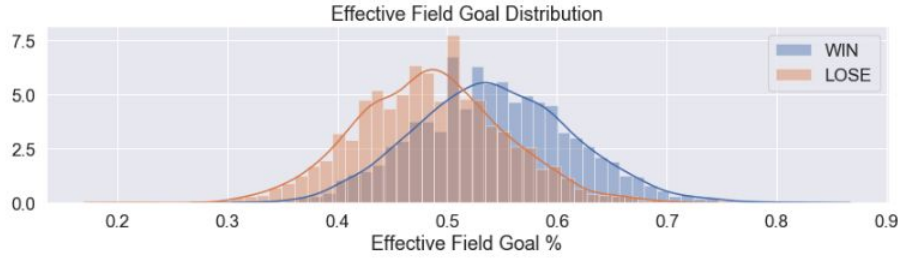
Home Team won 22.2% of the time when losing at the end of the third.

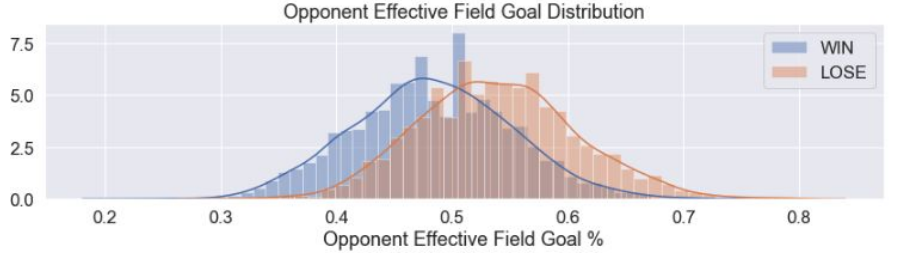
Shooting

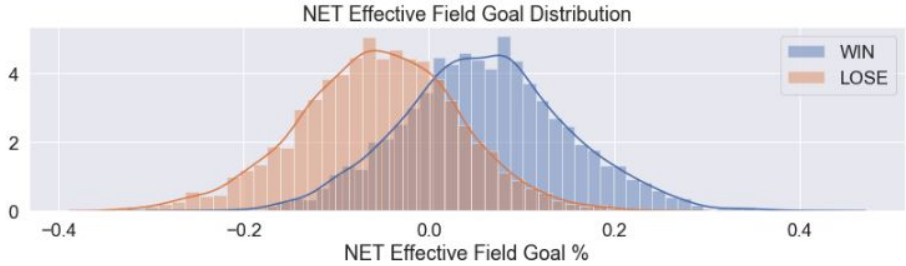
Key Takeaways:

Shooting isn't everything

NET Pos EFG % != Guaranteed Win
NET Neg EFG % != Guaranteed Loss







Model Performance: 82% Accuracy Score

TRUE POSITIVE (Pred: W, Act: W)	• 87% Correctly Predicted
FALSE NEGATIVE (Pred: L, Act: W)	• 13% Incorrectly Predicted (Type II Error)
TRUE NEGATIVE (Pred: L, Act: L)	• 75% Correctly Predicted
FALSE POSITIVE (Pred: W, Act: L)	• 25% Incorrectly Predicted (Type I Error)

Key Takeaways

#1 CURRENT +/-	• The 3rd quarter +/- is the most determining metric - *should be intuitive*
#2 SHOOTING & PROTECTING	• Shooting % Tier II Metric, followed by Turnover % (+1 TOVs => You -1 Chance, Opp. +1 Chance)
#3 OTHER GAME STATS	• Surprisingly, most other team statistics have very little impact (BLK, AST, FREE THROW RATE…)
FUTURE WORK	• Model that only predicts when 95% confident • Incorporate other popular advanced metrics

Any Questions?

THANKS :)