# Can we determine which NFL games to bet on by predicting the likelihood of an underdog team covering the spread?

Lavonne Hoang, Peter Holmes, and Jackson Hubbard

## Certain game conditions help the underdog cover the spread

Given our experience watching, playing, and betting on football games, we know that betting on certain teams are more profitable and further, that some game time conditions help underdog teams more so than favorite teams, and thus, help underdogs cover the spread. As a result, bettors can gain an edge over Vegas by placing bets against the spread informed by these factors. We aimed to quantify the effects of various factors on the predicted probability of the underdog team covering the spread in a given NFL game. The insights of our model can be used to help potential bettors make more confident and profitable bets.

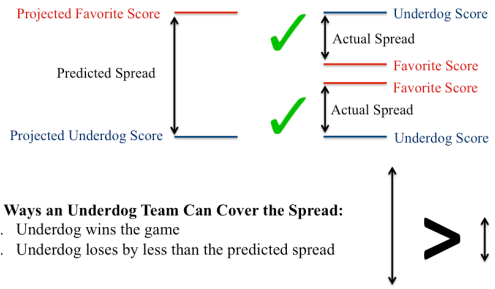
## We tested the effects of various game-time conditions

Our training data has 3,602 observations and includes information from the 2003-2016 season while our testing data has 255 observations from the 2017 season. After adjusting to only include significant variables, our final model predicts the probability of the underdog covering the spread using:

- **Precipitation:** indicator variable of whether there was precipitation when the game was played outdoors
- **Favorite in Best 3:** whether the favorite team was among the best 3 teams covering the spread (highest percentage of covering the spread)
- **Favorite in Worst 3:** whether the favored team was among the worst 3 teams covering the spread (lowest percentage of covering the spread)
- **Underdog in Best 3:** whether the underdog team was among the best 3 teams covering the spread
- **Underdog in Worst 3:** whether the underdog team was among the worst 3 teams covering the spread
- **Vegas Spread:** a continuous variable of the predicted spread by Vegas Sportsbooks
- Other variables tested for- Stadium, Division game, playoff game, Temperature

### Definitions

- **Underdog team:** The team projected to lose the game by Vegas sportsbooks
- **Favorite team:** The team projected to win the game by Vegas sportsbooks
- **Spread:** The amount of points the favorite team is projected to win by



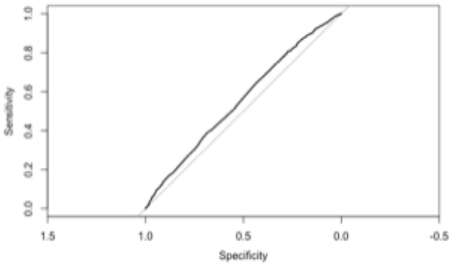

**2 Ways an Underdog Team Can Cover the Spread:**

1. Underdog wins the game
2. Underdog loses by less than the predicted spread

## Money Making Model

The area under the ROC curve is lower and misclassification rate is higher than what is normally seen in predictive models, however, we firmly believe that our model outputs great results given the context of our problem. Given how the Vegas spread is set so that 50% of bettors lose and 50% of bettors win, using our model gives potential bettors a slight edge over Vegas. In addition, predicting NFL games is not easy, as there are so many unquantifiable and unpredictable factors that go into winning a game. Thus, as a sports bettor simply being right as often as you are wrong is a great result. Our model does this and more as it is right about 55-60% of the time. This means that using our model will win a sports bettor money in the long run.

## Model used to correctly predict ~ 55% of games in training set




**Correct Prediction 54.06% of time**

## But if you bet intelligently… you can win up to ~ 65% of games bet on

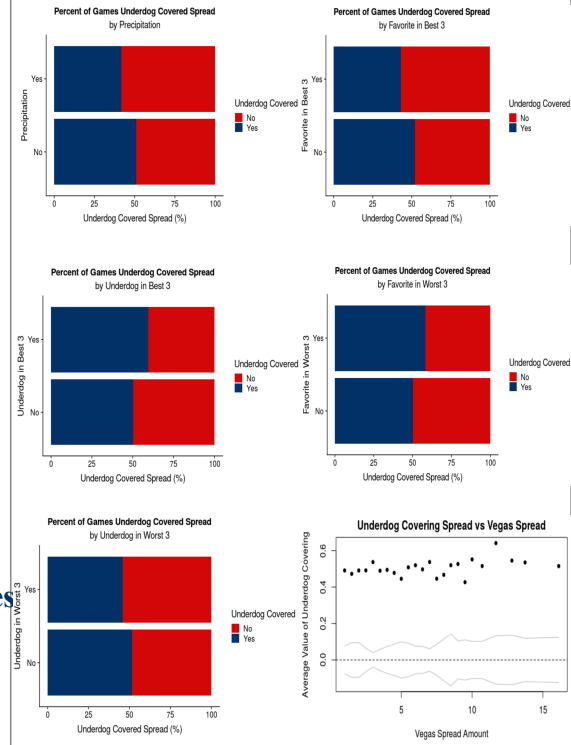
By interpreting the predicted probability of the underdog covering the spread, a bettor can use this information to bet on either the underdog covering the spread or on the favorite winning by more than the spread. For example, a conservative bettor may only bet on games with a predicted probability higher than a certain threshold (bet on the underdog) or lower than a certain threshold (bet on the favorite). Here, we simulated a conservative bettor placing bets on the underdog covering the spread if the predicted probability was >60% or on the favorite winning by more than the spread if the predicted probability was <40%.

**Betting $100 on each of the 84 games in 2017 that fit this criteria**

**Profited Us $1,000 ***

| | Estimate | P-Value |
|---|---|---|
| **Intercept** | -0.0361 | 0.59031 |
| **Precipitation** | -0.3614 | 0.04195 |
| **Vegas Spread** | 0.0247 | 0.01572 |
| **Favorite in Best 3?** | -0.3987 | 0.00004 |
| **Underdog in Best 3?** | 0.4151 | 0.00770 |
| **Favorite in Worst 3?** | 0.3440 | 0.03130 |
| **Underdog in Worst 3?** | -0.2686 | 0.00771 |
| **Underdog in Best 3? × Favorite in Worst 3?** | -2.1788 | 0.05583 |

## Gauging the effects of the explanatory variables




## Limitations to the model

The success rate of our model varies with the current week/season because of the amount of factors that go into winning a football game. Further variation comes from change within teams as players are injured, traded, and other similar unpredictable changes within teams. Additionally, the success of our model depends on an external factor, the Vegas spread, which essentially sets an upper threshold on the success rate. Our model is also limited by the fact that we determine the best/worst teams against the spread by using data from the prior season, but this is unreliable since teams change season to season. However, all of this is expected, as bettors face these same challenges, which is why betting on sports is so difficult.

* Bets not actually made. Data taken from testing data set (not included in model). Winnings adjusted for losses