Predicting Success on Third Downs in NFL Football

Problem statement

Third down and long is the toughest situation for any offensive coordinator in the NFL. – Matthew Stafford

Success in the NFL often comes down to consistent achievement in key but common situations. One of the most important tactical elements in NFL football is being able to convert a third down into a first down (or better). Several analytic studies have explored strategic considerations on fourth downs, but we are not aware of studies that have addressed success on third downs. In most third down plays (but not all), a team's offense tries to achieve a first-down conversion; if they fail, their options on fourth down are limited and much of the time the team will punt the ball.

The goal of this project is to obtain play-by-play NFL data to examine, predict, and summarize success on third downs. Situational features of the game are certain to play an important role: The number of yards short of the first down marker, where is the offense on the field, current score of the game, time remaining in the game, identity of key players such as the quarterback, running back, and so on. What types of plays are run are what are their success probabilities? What advice can you give coaches based on your conclusions?

Key Challenges: Scraping data, wrangling data, visualization, statistical modeling, machine learning.

Important milestone:

Data resources

You will be required to collect *play-by-play* data for NFL games. Some resources for play-by-play data include

- https://www.pro-football-reference.com/boxscores/201709100was.htm#all_pbp

- http://www.espn.com/nfl/playbyplay?gameId=400951760

- https://www.kaggle.com/maxhorowitz/nflplaybyplay2009to2016/data

It may help to use the `nflscrapR` R package. Details of the package may be found at https://github.com/maksimhorowitz/nflscrapR. A short example of its use can be found at https://www.r-bloggers.com/nfl-series/.

References

1. Berri, D. J., & Burke, B. (2012). Measuring productivity of NFL players. In The economics of the National Football League (pp. 137-158). Springer New York.

2. Fokoue, E., & Foehrenbach, D. (2013). A Statistical Data Mining Approach to Determining the Factors that Distinguish Championship Caliber Teams in the National Football League.

3. Onwuegbuzie, A. J. (2000). Is Defense or Offense More Important for Professional Football Teams? A Replication Study Using Data from the 1998-1999 Regular Football Season. Perceptual and motor skills, 90(2), 640-648.

4. Clevenson, M. L., & Wright, J. (2009). Go For It: What to consider when making fourth-down decisions. Chance, 22(1), 34-41.