NFL Capstone Proposal

Drafting a quarterback in the NFL draft is a notoriously difficult thing to do. The QB who excelled in college does not always translate that success upon entering the NFL. There are some guys who were “ok” in college that go on to become all time greats. Is there a way to make this seemingly unpredictable proposition more predictable? Imagine being an NFL General Manager. Wouldn’t it be handy to be able to predict the success of any given quarterback on your draft board with high confidence? This has been on the minds of NFL fans/GM’s/coaches for as long as the draft has existed. That’s the exact problem that I am setting out to solve with my first capstone with Springboard.

Client/Problem

To put it simply, as alluded to above, the NFL overall has a predictability problem when it comes to drafting college players. Year to year, the process of selecting a great QB in the NFL draft seems to be a coin flip. If a given NFL team is in desperation mode, whiffing on the draft might cost the coach or GM their jobs. With my capstone, I want to be able to offer the reader insights as to ***who*** some of these future greats might be. But I would also like to add another layer to this problem. There’s a real possibility that there may not be any discernible correlation between any physical/performance-based variables and a future NFL star whose draft stock is underrated, as is what I plan to explore with this project. There are plenty of instances every year in which a player is highly-touted heading into the draft, only to become a bust and vice versa. I want to explore **why that is.** How is it that a team like the Cleveland Browns (for example) can miss out on quality QBs when they have had prime draft positioning over the last few years? Why is there so much randomness and chance in the NFL draft in general? This will help to add a ‘real-world’ dimension to this capstone that the reader can hopefully see some value in (Even if they aren’t an NFL GM).

Data

For this project, I will be looking at NFL Draft and College Football Statistics data from 1970 - 2012. The NFL Draft dataset has a ton of statistics related to NFL career game performance for every player including passing yards, completions, sacks, tackles, etc (along with draft position). This will be useful for evaluating the long-term success of a player and whether or not they can be labeled a “franchise player” (A player who is often considered to be the best player on a team, or who is a younger player with the potential for a team to “build” around him). The College Football Statistics datasets gives information for every college football player and their statistics throughout their college careers. My idea is, by taking into account a response variable (Career Approximate Value) from the NFL Draft dataset along with college statistics, would there be a way for me to narrow down who the best available quarterbacks are in any upcoming draft?

Approach

My mentor, Devin, was able to share a project that a colleague of his made. In it, his colleague has a model that predicts whether a given quarterback will be a “bust” or “success”, which is very similar to what I would like to do. Ultimately, the goal for this project is to look at previous QBs who became successful and see if there might be some correlation between their success and some other variables. The simple idea is to predict who the best quarterbacks are in the upcoming draft by using the college statistics and respective career Approximate Value ratings for players that came before them. Part of my approach will also hinge on what I find during my exploratory data analysis.

Deliverables

This project could be presented in a variety of ways, but for now, I’m thinking that a Jupyter Notebook might be the best option for me. Although it would be neat to be able to present it as an interactive project - something that a reader could engage with.