Predictive Analytics as a Football Coaching Aid  
Project Proposal

John Rivas

Chris Bowes

*CS 6302 Predictive Analytics*

*University of Houston-Downtown*

[John@corporatemanaging.com](mailto:John@corporatemanaging.com)

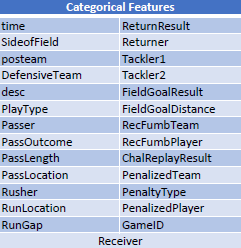
[bowes.chris@gmail.com](mailto:bowes.chris@gmail.com)

1. Abstract

There is an ever increasing reliance on data for decision making in the sports arena. Long term analysis on player development and performance projections have mostly dominated sports data analytics. But more recently, that data has begun to work its way into short term, more tactical decision making, instead of just the long term strategic decision making. In baseball, the regular use of the defensive shifts based on the batter’s tendencies is one primary tactical example. In this project, we aim to bring data driven analytics to football play calling to determine the most efficient/optimum choice.

1. Data Sets

The data to be used in this project will be both past/historical NFL game data, as well as current in-season data as games are played. This data consists of all game actions recorded by the NFL. For offensive players, depending on position, this includes things has how many carries, receptions, pass attempts and completions, yards gained or lost, how often a player is targeted. For more general team data collected, this will include attributes such as what type of play (run/pass), in which direction the play was run (in relation to the offensive team), the game time the play was run among other items. Most of the data would be numerical, the categorical items would include, play types, play direction as shown in table.

****

We would have several options for response variables. The tool would be designed to provide the best value play option given certain criteria, be that time of game, field position, defensive tendencies and offensive strengths. The best value will be defined as the type of play, the direction of the play that will yield the most positive yards gained for the team.

1. Challenges and Goals

In this project we aim to produce a tool that will produce a variety of results. While the predicted result of the play will be a regression problem, in predicting the amount of yards gained or lost, there will be a classification element to it as well, in that the tool must also decide from a discrete categorization of play types and directions.

One of the biggest challenges that will be faced is attempting to quantify the interaction between the offense and the defensive teams. Offensive stats are the easiest quantified, and defensive stats will mostly be defined in relation to the offense they were playing against at the time.

1. Methods and Tools

We will be using techniques that were first described in DeepMind’s AlphaGo program which was developed to play the game Go against human opponents. In their system, they used multiple nueral networks that worked in conjunction to decide on the next move. This combination of what they describe as a policy network, and a value network allowed AlphaGo to make decisions that were both strategic and tactical, while avoiding the tree search issues that plagued other attempts. Realizing that Go, shares many qualities in abstract, as the problem we are attempting to solve, we will implement a multi-network architecture to produce our results.

This data was acquired using the R package “nflscrapR” which allows the use of NFL JSON API data to perform detailed analysis on the games, players,and teams. We will be using python/TensorFlow to implement our networks.