

Fantasy Football Analysis

Ryan Carney and Blake Washburn
rjcarne (C01203776) and bwashbu (C89257841)

Our team has decided to pursue data analysis within the field of fantasy football because we understand the context behind the data and it is an area we both find interesting. For some background information on fantasy football, it is a simulated football season where players select professional football players from the National Football League and their weekly statistics in real life games determine whether or not fantasy football players win or lose. There is a ton of strategy involved in selecting your players each season and we wanted to create an application that we can apply to our own lives. To do this, we are going to create an application that performs a few machine learning techniques. Because so much strategy is involved in fantasy football, we can use our model to answer specific questions about different positions, such as whether it is better to draft a kicker from a good team or a bad team because they have inherently different experiences but our goal is to make data-based decisions instead of using intuition.

We have not fully decided on which machine learning methods we will use for analysis but based on the data and the problem at hand, we will probably end up using a predictive regression model or a classification model where we group players into tiers to evaluate their expected fantasy football value. Because football is such an unpredictable game, we will use all data points that we deem relevant to influence our model. K-means could be the route we chose because we can use a tier system since the most important statistic in terms of fantasy football is the points they earn towards a victory.

In order to evaluate our machine learning model, we plan to focus on specific positions and see how accurate our model is at predicting their fantasy point output. We can use some sort of reconstruction error to refine our model. One issue that we will need to be aware of is that because players are in the National Football League for many years, a player's statistics from each year need to be independent from other years instead of using a career total because the variables around them constantly change.