

NFL Play Predictor

Chapman Beard

Description: I am a very big sports fan, but the NFL is by far my favorite league to watch and follow. I watch games for hours each week, and keep up with all the latest trades and news throughout the week and offseason. I played football for many years and the idea of using machine learning to be able to predict certain plays and things that will happen in the game is very interesting to me. I have found a dataset with about 100 different features with every play from every NFL regular season game from the years 2009-2016. Not all of these features will be used, but I would like to train a model to predict the next play in an NFL game. Not only whether it was a run or pass play, but which gap the run was, or the distance of the pass, and so forth.

<https://www.kaggle.com/datasets/maxhorowitz/nflplaybyplay2009to2016>

Data: Many useful features in this dataset may include which team is home vs away, time on the clock, timeouts, game score, etc. I think it would be fairly easy to get a model to predict whether a play would be a run or pass, but I am curious to see how specific we can get the model to predict with a certain degree of confidence. This means the data would take the following form:

PlayType	Qtr	Down	Time	yardln	ydsnet	PassLength	RushLocation	etc.
Rush	1	2	14:16	47	2	N/A	Right	
Pass	4	3	3:43	20	13	10	N/A	

To get this data, I already have the data set from kaggle with around 100 features and around 360,000 different plays. This data, however, would need to be cleaned and prepared for use. Some of the features like the different player id's and tackler would not be necessary and could be omitted. In addition to this, we can probably throw out all kicking plays as this will not be applicable to our goal we are trying to achieve.