

Project 4 Proposal:

Team Members: Carson Young, Joe Uva, Julie Johnson

Overall Objective:

The objective of our project is to develop a machine learning model that predicts match outcomes in the Premier League based on historical data, match-ups, and other relevant variables. By scraping data from websites using Python, we will gather comprehensive Premier League data spanning multiple seasons. The collected data will be used to train our machine learning model to hopefully allow it to make accurate predictions for upcoming matches. We will also employ data visualization techniques in Tableau to present our findings and insights in a visually engaging manner.

Data Collection:

We will leverage Python's BeautifulSoup library to extract relevant data from Premier League websites by going through the CSS and HTML code that makes up the site in order to get all relevant information. This will include information such as match outcomes, offensive and defensive stats and other variables we deem influential in deciding a match.

Data Preprocessing:

Once the data is gathered we will clean and transform it into a suitable format for machine learning analysis using the Pandas library within Python.

Machine Learning Model Development:

Using the preprocessed data, we will build a predictive model capable of forecasting match outcomes. We will explore various machine learning algorithms in order to determine which algorithm achieves our team's goals based on the training and learning data presented to it. The developed model will be evaluated for accuracy, precision, and F1-score.

Visualization in Tableau:

In order to display our findings and conclusions while working with our machine learning model predictor, we will implement the use of Tableau in order to visualize the models overall success and accuracy as a match outcome predictor.