PGA Score Predictor Design Document

By: Nick Ospelt; Created: 9/10/2024; Updated: 9/10/2024

Problem Statement

Golf is an exact sport. The difference between positions is one stroke throughout 72 holes. Also, so many factors external to how the player is playing determine their score. For example, the course conditions and the weather the players are playing under. Is there a way to predict how a PGA tour player would perform on any given day, taking all of these factors into account?

Solution

Using historical weather data, player statistics, course information, and tournament results develop a model that predicts a score, the number of strokes it would take a certain player to finish 18 holes. Once that model is developed, productionize the model in a way that users can interact with the model and leverage the predictions the model makes.

Data

Weather (Open Meteo API):

Average Daily Temperature (*F), Daily Precipitation Sum (in.), Windspeeds (mph)

This information should correlate to the day and location where the round was played.

Player Statistics (PGA Tour Website):

Greens In Regulation (%), Birdie to Bogey Ratio, Fairways Hit (%), Total Driving Distance (yds.), Actual Scoring Average (Strokes), Scrambling (%), Putts Per Round (Strokes), Three Putts Per Roud

This information should be the player averages going into a specific tournament.

Tournament Results (ESPN Website):

Course Par (Strokes), Course Distance (Yds.), Course Rating (Strokes), Player Score (Strokes)

Model

The model will try to predict the player's score, which is in the form of number of strokes. This value generally ranges between 60-90 for pros. Therefore, this will be a regression problem.

Possible Models:

1. Linear Regression
2. Decision Tree/Random Forrest
3. Neural Network
4. Support Vector Machines

Possible Performance Metrics:

1. Mean Squared Error
2. R^2 Error

Model Production

Save the model and containerize it in a docker image.
Create a rest API that can be used to interact with the model from the end product

End Product

A web application where the users can input weather, course, and player data and get a prediction for the round.

Possible Features:

1. Pre-inputted player data (Select a certain player without having to find the data themselves)
2. Pre-inputted Course data (Select a certain course without having to find the data themselves)
3. Way to generate multiple round predictions at one time (Same player and course, but different weather conditions)

System Diagram