Tierney McBride

tierney6@pdx.edu

Project name: Are We Gonna Win the League?

Project area: Statistics, Sports Statistics, Using Concurrency for Large Calculations

Description of project:

  After nearly 20 years of being a casual fan of strictly international soccer, I have found myself in the last couple of years becoming a deeply dedicated fan of a team in the English Premier League. As an American who grew up in a baseball household, I often find myself longing for the soccer equivalent of the "magic number" – the statistic calculated to determine how many games a team would need to win to clinch a spot in the MLB playoffs. In the Premier League, there are no playoffs, but the spot in which you finish in the standings does determine whether or not you qualify for bigger tournaments. Add to the picture the fact that your standing is determined by the points you earn from either winning or drawing a game, and determining whether your team still has a chance at any given place at any given time in the season becomes overwhelming relatively quickly in most circumstances. Hence, this program.

  The goal is to create a web program to determine whether or not a given (user-provided) team can reach a given spot in the final standings of the English Premier League, given current standings. Using an API to fetch current standings data and team-specific performance data, the program will run a Monte Carlo analysis to determine the probability of every possible remaining outcome. This will leverage Rust's web and concurrency capabilities.

  MVP: Given a user-provided team name, the program will determine the probability of said team achieving 1st place in the final standings of the English Premier League.

  Additional features: Calculate simple probability of a finish in any given place; display possible avenues for achieving said finish; identify which teams needed to beat in order to achieve said finish

Possible issues: Identifying an API/data source that has the desired data for the model/having to use multiple data sources; that it's been too long since I've taken a math class covering probability and statistics, so I may be calculating something other than I think I am, but in that case, I am willing to alter the math or the description of the project, depending on time remaining to finish the assignment

Git repo: github.com/beansighe/arewegonnawintheleague

Sources:

Model for Monte Carlo simulations will be informed by the model in this paper:
https://math.uchicago.edu/~may/REU2022/REUPapers/Akuzawa.pdf