Assessing the Probability of the Soccer Match Outcome

- Team members
- The project description
- Proposed Work
- Prior Work
- Evaluation
- Datasets and List of Tools

Team Members

- Arman Aydemir
- Kamiar Coffey
- Shanqing Yu
- Xinyuan Zhao

The Project Description

The Project Description

- The goal of this project is primarily to create a prediction model which can assess the probabilities of the soccer match outcome. The second goal is to devise a betting strategy which in long can win when betting at bookmakers.

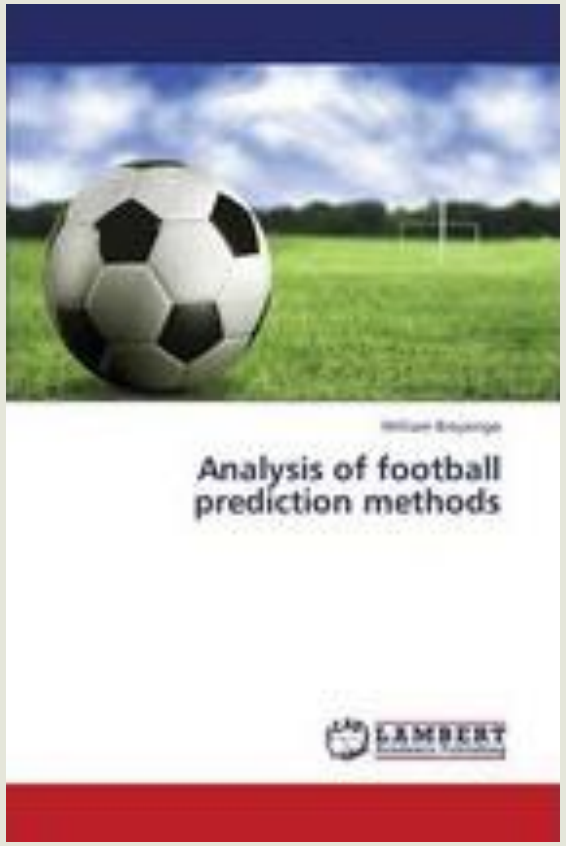
Questions

- Are there significant betting opportunities for soccer betting?
- Is it possible to build a successful model for predictions of the results for soccer matches?

Prior work

- Prior work base on Dixon-coles model.
- http://web.math.ku.dk/~rolf/teaching/thesis/DixonColes.pdf

- Dixon-Coles Model is used to find a betting strategy based on the full-time scores & the difference of goals.
- Rather than just use the goals of two teams, we hope we can get more attributes involved to find a more accurate model.



Proposed Work

Data Cleaning

- Only the information about the numbers of the goal scored, when and where the score happens is relevant, other information such as the referee's name, the number of yellow or red cards will be discarded.

Data Preprocessing

- The ratio of frequencies of home wins,draws and away wins will be used to determine the home advantage. The dependency between home and away score will also be checked.

Data Integration

- A large amount of data results from different seasons will be integrated.

Evaluation

- The model for determining the probabilities of each soccer match outcome will be developed. It will be compared with the bookmaker's assessment with regards to setting odds. The performance of the proposed model and betting strategies performed on actual odds will be examined

Datasets and Tools

- List of datasets to use

  Dataset of English premier league matches

  https://datahub.io/sports-data/english-premier-league

- List of Tools

  Python and Scrapy (A web scraper to scrap online data)