School of Electronic Engineering and Computer Science

Queen Mary University of London

MSc PROJECT DEFINITION

2021-22

Project Title: Predicting the final seed of National Basketball Association teams, an Elo based approach

Supervisor: Mr Hazar Emre Tez

Student name: Thomas Abraham

Student email: t.abraham@se21.qmul.ac.uk

PROJECT AIMS:

The aim of the project will be to create a model that can accurately predict each of the 15 seeds in both conferences for an upcoming season. It can be tested by applying it to previous seasons where the outcome is already known. In this project, an Elo-based approach is employed to model an individual basketball player's strength based on the plus-minus score of the player and obtain a rating. The plus-minus score is a useful metric because it quantifies the contribution of a player like good defense and setting up screens which are not reflected by metrics that are primarily based on scoring. The individual player ratings are combined to obtain a team rating. Team ratings are compared pairwise to obtain the probability of a win by each of the teams during the course of a season. The rating system is validated by running them over real-life data from past NBA seasons.

PROJECT OBJECTIVES:

- Understand and be able to apply a modified version of the Elo algorithm and the K factor.
- Update the algorithm to the desired specifications.
- Identify the required size of the data-sets. (If more data is required then more seasons will have to be considered and vice-versa)
- Perform numerical simulations

METHODOLOGY:

- Completely understand the Elo algorithm and its methodology.
- Implement changes to the algorithm in order for it to better suit the project.
- Scrape the real-life data available on the NBA Advanced statistics website.
- Perform simulations using the data gathered so far.
- Identify if the aim was achieved by testing it on previous seasons.

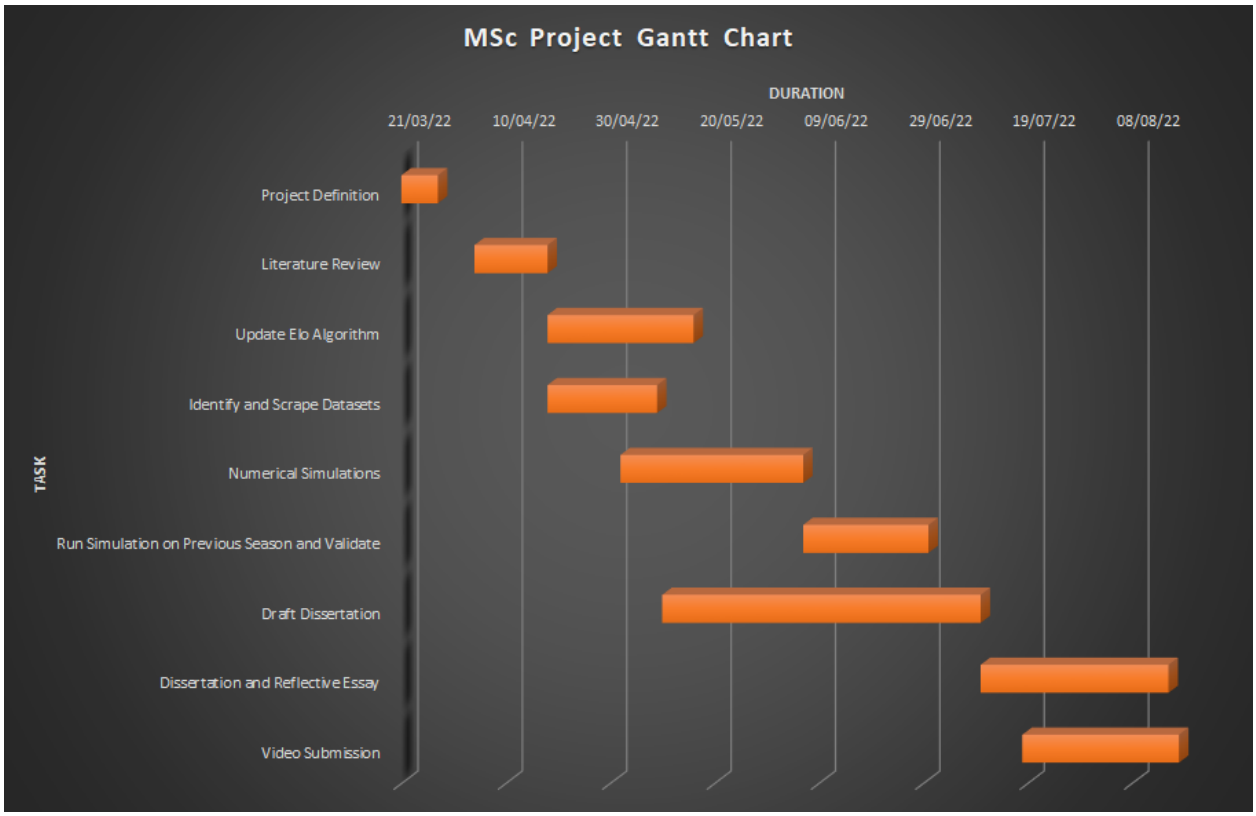
PROJECT MILESTONES

- Completing detailed literature review
- Understanding and updating the Elo Algorithm
- Scraping datasets
- Testing the updated elo algorithm
- Numerical simulations on previous seasons are successful
- Draft of final dissertation is completed

REQUIRED KNOWLEDGE/ SKILLS/TOOLS/RESOURCES:

- Python for data scraping
- Jupyter Notebook will be used for mathematical modeling

TIME PLAN


MSc Project Gantt Chart
DURATION
21/03/22
10/04/22
30/04/22
20/05/22
09/06/22
29/06/22
19/07/22
08/08/22
TASK
Project Definition
Literature Review
Update Elo Algorithm
Identify and Scrape Datasets
Numerical Simulations
Run Simulation on Previous Season and Validate
Draft Dissertation
Dissertation and Reflective Essay
Video Submission