Archit Bose
Nitish Padavala
Gabe Hyun

First Draft of Group 9 Project Pitch/Summary

Pitch:

The big picture of this project is to use Machine Learning on NBA Player data in order to accurately predict which NBA team is likely to win a certain basketball game. In order to gather this data, data will be collected from Basketball-Reference.com's dataset which accurately records player data day by day. Using this data, models can be generated to predict the likelihood of winning games by combining the statistics of individual players to compute a score for any given team on a certain game day and compare that score with that of the opposing team. A reinforcement step will be added to learn from new data based on who the actual winner of the game is on a certain day to better predict future games.

Specific Aims:

What models and features would be most helpful in creating an accurate prediction?
How accurately can we predict the winner of a given NBA game?
Is there any particular statistic that clearly indicates the likelihood to win a given NBA game?
Do better individual player performances correlate to better team performances?

Timeline:

Week 4:
Scrape the data from https://www.basketball-reference.com
Week 5:
Cleanse the dataset to make it suitable for ML training
Week 6:
Conduct an EDA and identify predictive features
Week 7:
Create and train a neural network on the dataset
Week 8:
Test model and evaluate success/failure of the model using criteria such as accuracy, recall, f1 score, etc..
Week 9:
Improve model accuracy and tune hyperparameters
Week 10:
Create presentation on Google Slides and write supplementary report if needed

Team Outline:

- Archit Bose
  - Generate equations to predict player statistics in relation to overall team score
  - Analyze model and evaluate accuracy, precision, etc. (validation)
  - Test model and add learning feature based on new game data
- Nitish Padavala
  - Decide on features used in model
  - Train Model
- Gabe Hyun
  - Scrape data from Basketball Reference
  - Clean the data and process it into Team and Player DataFrames

Final Deliverable:

Our final deliverable will be in the form of a presentation and a supplementary report if needed