DeepNBA Final Presentation

Nikhil Sinha, Jason Witherspoon

Motivation

- Our aim is to predict the outcomes of NBA games
- Current prediction models use mostly high level, coarse data like win/loss records
- We are trying to flip this approach head and use super fine, raw box score data to project a win or loss
- Accuracy goal: >67% (Las Vegas sportsbook projection accuracy)

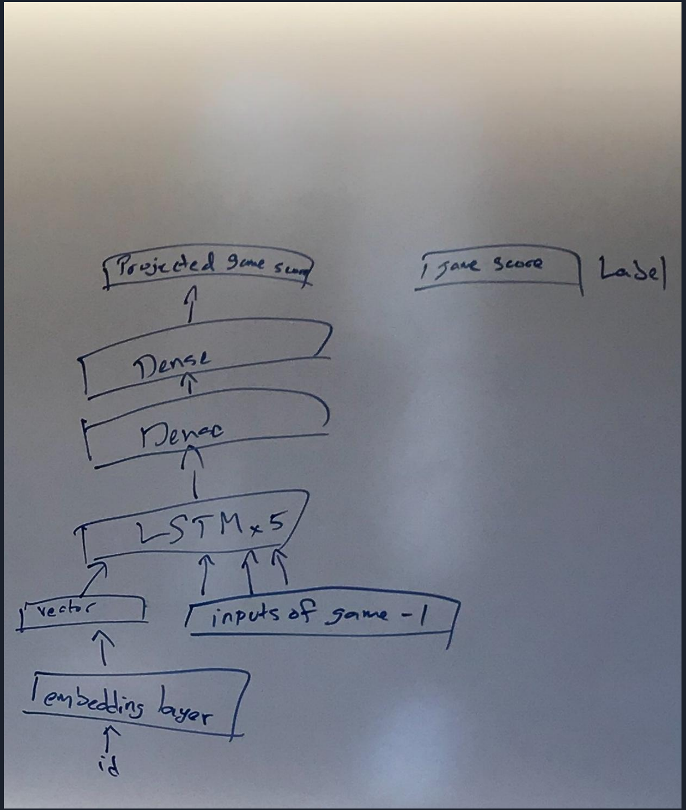
Model architecture

- Game data segmented into players, teams and peripheral data
- For each player, grab the past 5 games that they played in and feed into an RNN to project a game score (formula devised by John Hollinger at ESPN) to estimate individual box score performance
- For each team, grab the past 5 games that they played in and feed into an RNN to project a team score to estimate team valued offense and defense
- Pass individual player projections, team value projections and peripheral data (like distance travelled and W/L record) to a FFNN to project win or loss for the home team

Quick stats

- 13,660 games
- 27,320 individual team stat lines
- ~280,000 individual player stat lines

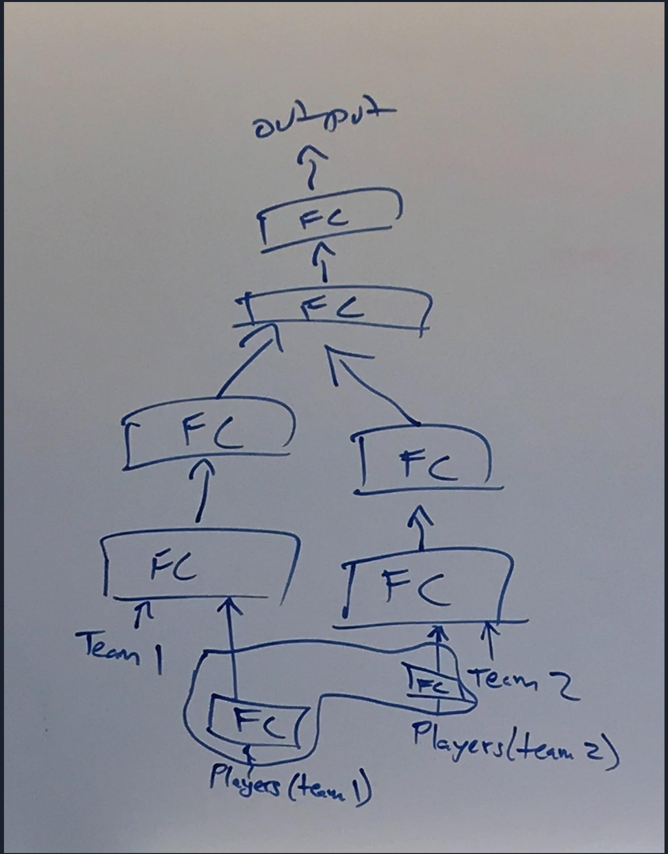
Player model/Team model



Stuff We Learned

- Choose the right loss function - MSE is not that good for things where you want to mitigate large incorrectness (like snapping a streak)
- Create metrics you care about
- Embedding layer for representing different players/teams

Prediction model



Prediction model specifics

- Trained using the predictions from our earlier models to be somewhat like end to end
- Concatenation of different segments - allow model to learn from the team first before learning from the different teams

Model Results

- Player model: ~80% predictions within 2 of intended - good!
- Team model: ~40% predictions within 1 of intended - not good
- Prediction model: ~55% accuracy - also not good