Meeting – 2/7/2025

* Discuss initial thoughts about modelling the number of runs scored per bowled ball
* Discuss alternative machine learning (neural network) model to improve computational efficiency while not losing any predictive ability
* Discuss import metrics to ascertain from the model output
  + Average number of runs scored per 100 balls for a given player, league year combination?
    - Incorporate grounds into calculation?
    - Starting from 1st batter in 1st innings for fair comparison?
    - Different starting place in batting order?
  + Correlation structure between player’s batting ability between different leagues
    - Batter rating for each league similar to the hockey paper?
  + Scatterplot with the probability of a six against the expected value of a ball for each batter for each league aggregated across all possible bowlers
    - Visualize hitters (4 and 6) vs. batters (1 and 2)
* Questions for Hassan about the data/cricket in general
  + If we want to incorporate information about the grounds into our study, how can we account for the changes in the boundary line between matches?
    - Ignore for now
  + How should I account for the SuperOver data? Should that be excluded?
    - Ignore
  + You mentioned that you want to take into account the handedness and pitching style of the bowler. What dataset in the package can I find this information?
    - Will get the data from Hassan
  + The information for the season is different for different leagues (i.e. some seasons take place over two calendar years and some take place over one calendar year). How should we account for this.
    - Change season to the calendar year in which it started.