Solution Sheet

1. Which model have you used for Total IPL 2020 Runs prediction for each player? Explain your model.

   - We used Gradient Boosting for regression. GB builds an additive model in a forward stage-wise fashion. Gradient boosting can be used for regression and classification problems. it allows for the optimization of arbitrary differentiable loss functions. In each stage a regression tree is fit on the negative gradient of the given loss function.
   - While implementing the model we used the value of n_estimators as 55. We did not use a bigger value to avoid over fitting, because of the small size of dataset. In large datasets a larger value results in better performance.

- For loss function we used 'huber'. Because it is a combination of 'ls'(least squares) and 'lad'(Least absolute deviation) which is a highly robust loss function solely based on order information of input variables.