**Solution Sheet**

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

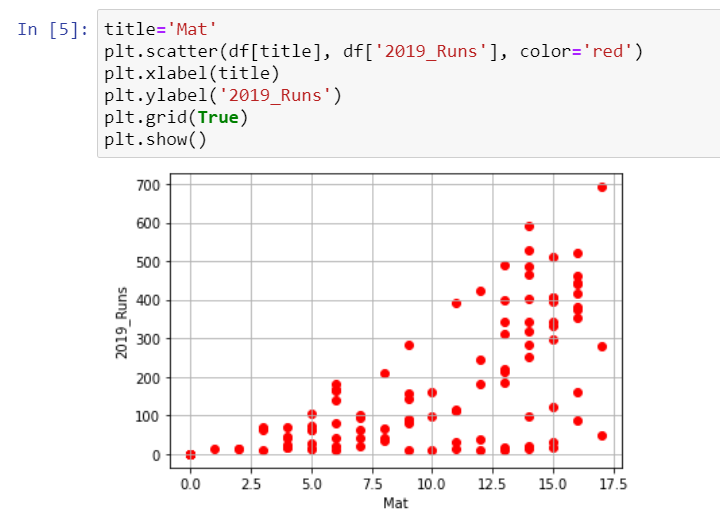
Model used: Multiple Linear Regression

Step 0: Clean the dataset of any Nan or special characters so that the data to be handled is completely numeric in nature.

Step 2: Set aside some rows from the given training set for finding the accuracy and testing the performance of our model. Here 10 rows are extracted for the purpose.

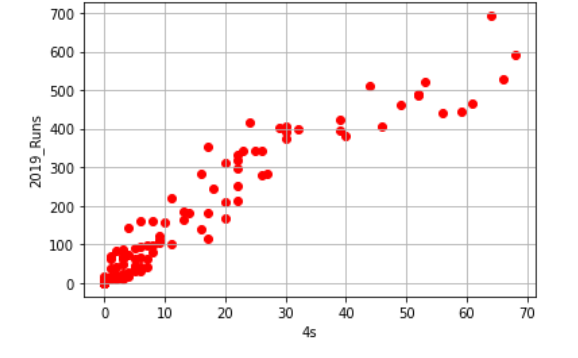
Step 2: Find out the relationship between the various input variables and output variables in the training dataset:

For example:



This shows as number of matches increase, there is an increase in the number of players scoring higher runs. However, there are players with low runs as well as matches increase.

Another example:



This shows a good correlation between total runs scored and total number of 4s hit in the tournament.

Similarly, find the linearity and choose the variable with good linearity and correlation for the next step.

Step 2:

The columns thus chosen are:

'2018\_Runs', 'BF', '4s'

Now find the linear coefficients of each of these variables with respect to the output (2020\_runs).

where are the coefficients and c is the y-intercepts

The coefficients obtained are:

[ 0.74553157 0.53010922 -1.35739413]

And the intercept c is:

-3.02215188573507

Step 3:

Use the intercept and the coefficients on the corresponding columns of the test data set and predict the total runs in the dataset we had set aside for measuring model accuracy.

Since the values are continuous in nature and not as classes, we cannot use the accuracy function in scikitlearn. Instead we use r2\_score for measuring the similarity between our predicted and original values.

We get the r2\_score to be: 0.8800014416812713

For the same data set, higher R-squared values represent smaller differences between the observed data and the fitted values.

Step 4:

Use our model on the test data set and predict the total runs to be scored in 2020.

Done!

Thank you Flipr Hackathon 7.0 for boosting my confidence in using ML for such prestigious hackathons.

Individual Submission by:

Ritu Ann Roy George

ritu\_b170106ec@nitc.ac.in