

ICC World Cup 2018 Player of the Tournament Prediction

Group ID - A2

Lab Group No - 2

Shihab Shahriar Ahmed - 14.02.04.029

Shehabul Alam Shatu - 14.02.04.030

1 Problem Description

The ICC World Cup will be held later in this year. Here, we try to predict the rating of the players. Using those ratings, we can easily predict about the next player of the tournament. As we are using supervised machine learning, we have to set all the ratings at the beginning so that we can check the accuracy of our prediction later.

2 Dataset

We have a custom made dataset named 'cricket.csv'. It has eight columns and sixty rows. The attributes of our dataset are 'Id', 'Name', 'Age', 'Match', 'AVG', 'SR', 'Team Point', 'RATE'. Of them, we have used six attributes as 'id' and 'Name' are valueless here. Out of six attributes, the first five are considered as feature and we set 'RATE' as our target. Every feature in this dataset is in numerical form. To make this dataset we have used <http://www.espnccricinfo.com/> and <http://www.cricbuzz.com/>

3 Model

We have used four different models here. They are:

- Linear Regression
- Decision Tree Regression
- Random Forest Regression
- KNN Regression

4 Comparison of Performance

Performance	Liner Regression	Decision Tree Regression	Random Forest Regression	KNN Regression
r2 score	-0.1663	0.3024	0.1849	0.1548
Mean Square Error	24,056.0207	15,313.4	15,252.4272	32547.96
Mean Absolute Error	127.5573	98.6	99.1456	199.76
Medium Absolute Error	26.5801	9.5	14.616	14.6

Table 1: Performance Matrics

5 Discussion

From the above discussion and performance matrices, we have seen that Decision Tree Regression performs better than the other three regression. As we have only sixty tuples in our dataset performance of the project is not as high as expected.