



CSE623: Machine Learning Theory and Practice

Report-3

Group 1

Name	Enr No.
Dhaivat Patel	AU2240022
Dhyey Patel	AU2240054
Tirthraj Raval	AU2240079
Anusha Jain	AU2240092
Sloka Thakkar	AU2240103

Work Done:

- Performed Exploratory Data Analysis (EDA) to detect the outliers, to handle missing values if any, to understand the relation between various features.
- Read through various research papers to finalise appropriate formulas to calculate the offensive and defensive rating and use them as a base for the regression analysis and offense, defense categorisation.
- Finalizing, logistic regression as the algorithm to predict the new player data whether she is offensive, defensive or all rounder from the k-means clustering algorithm output as parameter that forms the clusters.

Offensive rating:

- $ORTg = 100 * (PProd / TotPoss)$

$PProd = \text{Individual points}$ and $TotPoss = \text{Total Possession}$

Defensive rating:

- $DRtg = Team_Defensive_Rating + 0.2 * (100 * D_Pts_per_ScPoss * (1 - Stop\%) - Team_Defensive_Rating)$

Goals for next week:

- For predicting the optimal line up against a specific team we are going to discuss with the professor and the mentor.
- Implementing the process which the professor and the group will decide after the discussion as mentioned in the above point.

References:

D. Oliver, *Basketball on Paper: Rules and Tools for Performance Analysis*. S.l.: Potomac Books, 2020.