****

DEPARTMENT OF INFORMATICS

SCHOOL OF COMPUTER SCIENCE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES, BIDHOLI, DEHRADUN, UTTRAKHAND, INDIA

Synopsis

**Title:** Player Selection (IPL Team) using Clustering Algorithm

* **Introduction:**

IPL also known as Indian Premiere league launched in 2008 by BCCI (Board of Control for Cricket in India) is a professional Twenty20 cricket league in India. It is not only loved by fans from all over the world but it also plays a very significant role in India’s Economy. IPL consists of 8 different teams representing 8 different cities in India. It is also a platform for new and young talent to make their way into the Indian Cricket team. This year due to the COVID-19 pandemic it was postponed but since it holds so much importance that BCCI decided to organize the tournament in United Arab Emirates (UAE) in September 2020.

* **Problem Statement:**

In any game where teams are involved, selection of a good team and coordination of that team is a very important step. IPL has a similar format of picking your own team but they do through players’ auction where which player will get to play in which team gets decided and the highest bidder gets the player. Sometimes the player’s base price or starting price is in crores and when there is so much money involved then there is no scope for mistakes. That is where data analysis comes in handy , it helps us make a decision based on data and most of the times that is the best decision. The main objective of the project is IPL Player selection for a team by analyzing previous data. When you are searching for players you can select the parameters (on the basis of which they are to be clustered). Keeping in mind the result of clusters, the budget of the team and preference of players we can select a really good team at the time of team selection or players auction and hence the results and can be improved. The probability between match winning VS toss winning like it is displayed on the television during the telecast will also be shown.

* **How we will solve this problem?**

Predicting the events that may happen IPL follows the past data so we will first read the data from CSV file then preprocess it with the desired formats according to the needs. Effective use of File Handling and Array of Structures as the storage options and we will extract the data . Cleaning the data and making it ready for the analysis will be the first step. Then diving deep into the clustering concepts, we use K-MEANS clustering Algorithm to create clusters of the players considering both the bowler or batsman and the whole scenario will be based on similar characteristics such as bowling speed, runs scored and the maximum wickets taken. Analysing the clusters made we will focus on predicting how the factors will improve the game for players. Prime focus of selecting the cream players from the group's/cluster's will give us our dream team.We will also calculate the probability between match winning VS toss winning.

Project Created Under

**Mr. Deepak Kumar Sharma**

Created By

**Name –** Lakshay Vasuja **Name** – Lakshay Sharma **Name** – Divyansh Chandna

**Roll no –** R172218064 **Roll no** – R172218063 **Roll no** – R172218045

**Sapid –** 500067177 **Sapid** – 500068760 **Sapid** - 500068095

**Branch** – CSE-BigData “B2” CSE-BigData”B2” **Branch** - CSE-BigData”B1”