

**School of Computer Science**

**University of Petroleum and Energy Studies**

**Dehradun-248007**

**2018-22**

**MINOR PROJECT 1**

Title: Player Selection (IPL Team) using Clustering Algorithm

**Under the guidance of**

**Mr. Deepak Kumar Sharma**

Assistant Professor

Department of Informatics

**Submitted by**

LakshayVasuja , 500067177,R172218064

Lakshay Sharma , 500068760, R172218063

DivyanshChandna, 500068095, R172218045

MID Semester Progress Report

Title:Player Selection (IPL Team) using Clustering Algorithm

**Problem Statement**

There can be a number of situations where desired players cannot be picked during auction so a substitute player has to be picked.

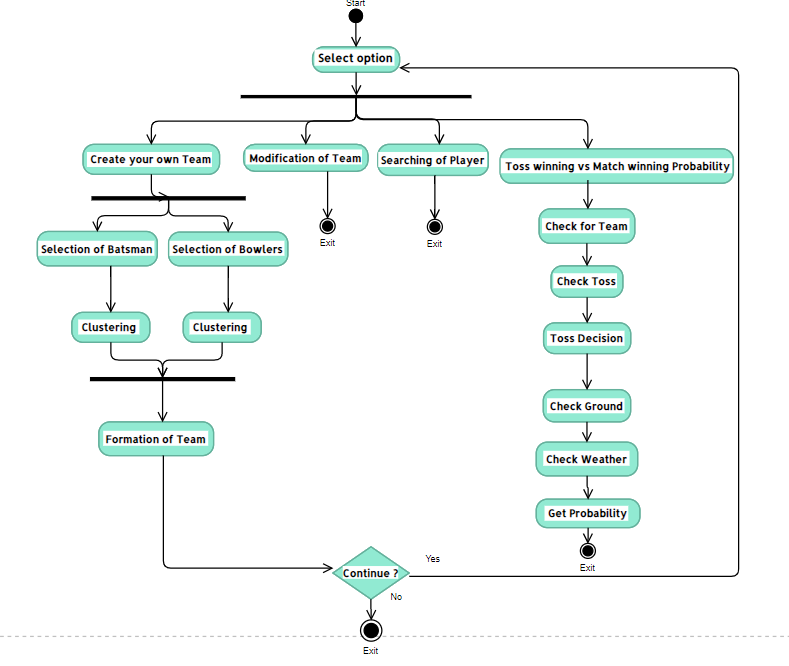
To make a good and versatile team with different properties is a hard task as team management have to keep track of all the players and bid on them.

Sometimes Captains cannot make a well informed decision on what to choose after winning the toss.

**Methodology**

1. Predicting the events that may happen IPL follows the past. Cleaning the data and making it ready for the analysis will be the first step.
2. We start with a first group of randomly selected players(centroids), which are used as the beginning points for every cluster.
3. Iterative calculations are performed to optimize the positions of the players in the clusters.
4. Now the clusters are formed the players are ready to be selected.
5. After selection of players, the dream team is formed.

**Flow Chart**

****

**Algorithm**

1. Choose the number of clusters(K) and obtain the data points
2. Place the centroids c\_1, c\_2, .....c\_k randomly
3. For each data point x(i):
   * + 1. find the nearest centroid(c\_1, c\_2 .. c\_k) using Euclidean distance
       2. assign the point to that cluster
4. For each cluster I = 1…k
   * + 1. new centroid = mean of all points assigned to that cluster
5. Repeat steps 4 and 5 until convergence or until the end of a fixed number of iterations.

**Implemented Code**

Functions:

1. **Random()**

The first step of k-means is carried out with the help of this function. It selects a random index. All the values at that random index are taken as the initial values and then the Euclidean distance from that point to others is calculated and an array of those distances is returned.

1. **fileWriting()**

Here 3 inbuilt functions of file handlingfwrite,fseek,fread are used.

fwrite- Creation of text file of selected batsman

fseek- Traversing of selected players in text file

fread- Reading of selected player

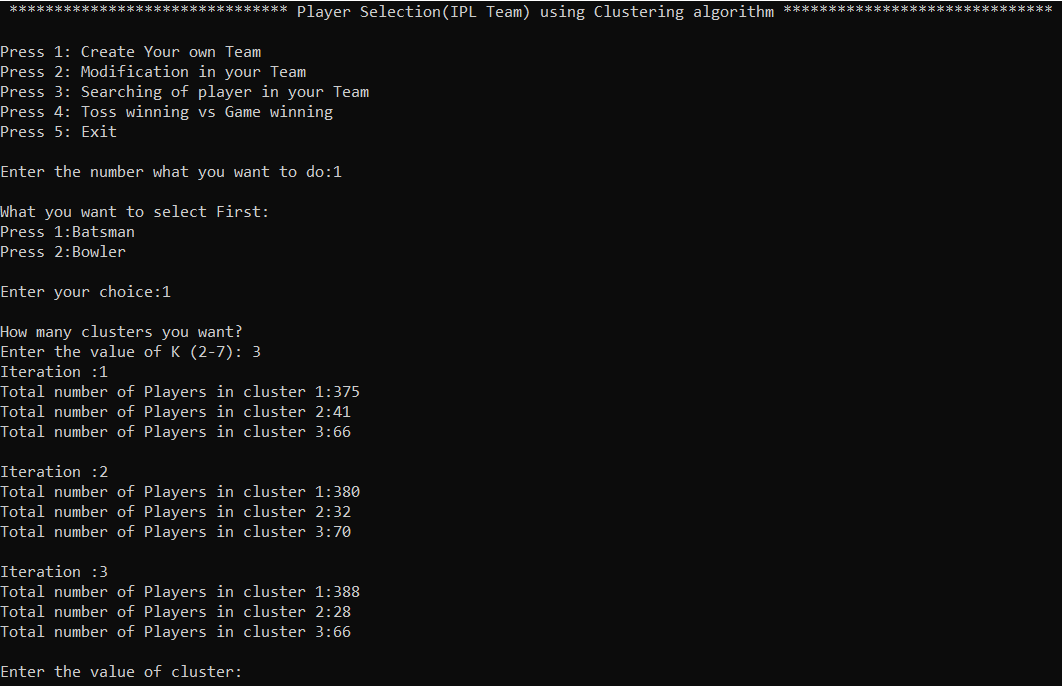
1. **getProbability()**

First of all it asks for a number of inputs like selecting the team, outcome of the toss, their playing conditions and then it will calculate the probability of the match winning.

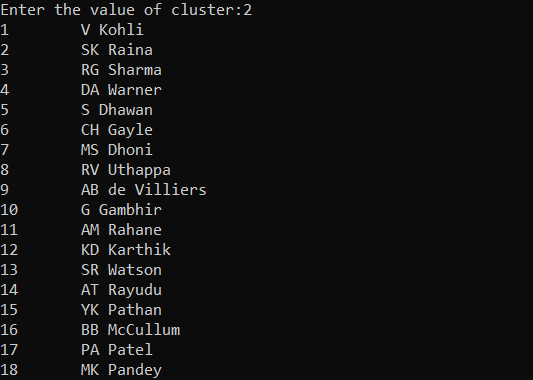
1. **displayProbability()**

This is an output function that displays the matches won, matches lost and total matches played and then if there is data to calculate the probability it will display the probability of match winning and match losing.

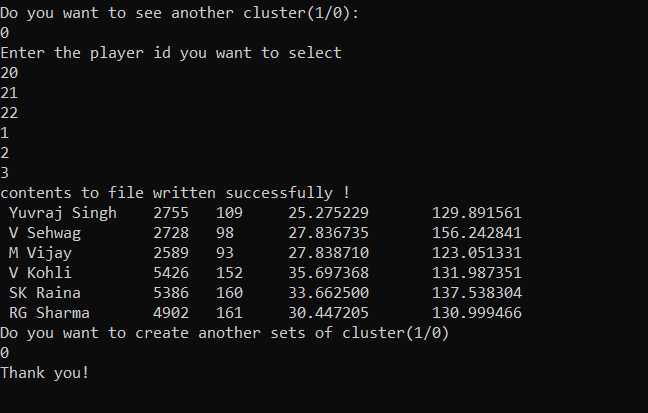
**Results**

****

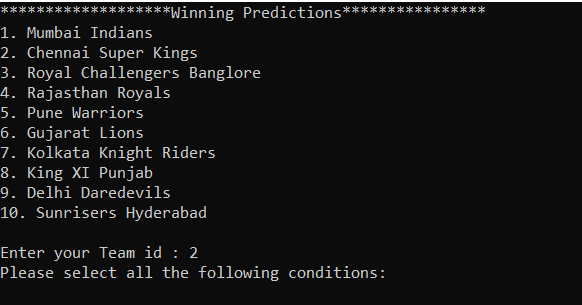
**List of cluster 2 Batsman**

****

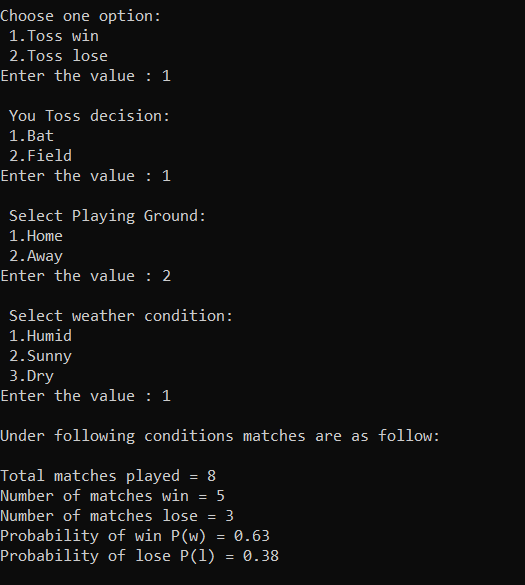
**Choose batsman according to choice and store in file**

****

**Selecting team for finding probability**

****

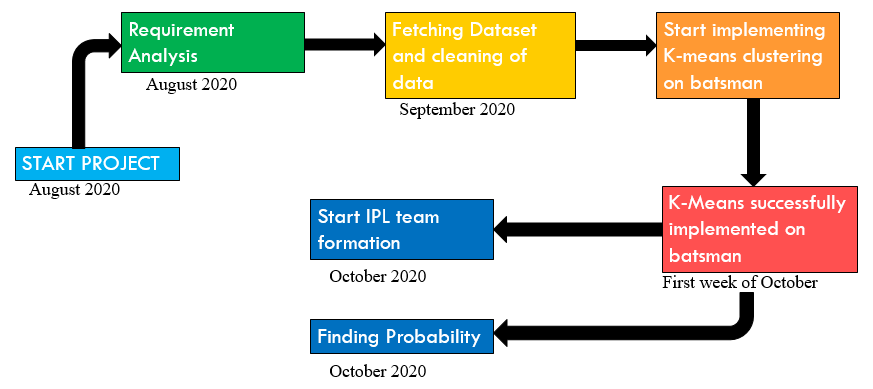
**Under selected conditions Finding Probability**

****

**Scope of project**

1. Selecting a team having best players comprises an important part of IPL auctions. We covered this scenario with the scope that there may arise situations where desired players are already booked, creating clusters of players having similar properties will help in choosing the substitute of the desired player.
2. In an IPL match, captains have an important responsibility of choosing field or bat after winning Toss, having a proper knowledge of pitch, weather conditions in which they have played will help captains to understand their hold over the match.
3. Every year many young players (say 300) present themselves in the auction, a few around (say 200) gets the chance to play. Teams are categorized having 16 players each, here arises the problem of grouping the data of all players such that we create a balanced team having good to average players considering the bidding amount.

**Work flow(Pert Chart)**

****

**Contribution of each member**

**LakshayVasuja**

1. Finding and cleaning dataset
2. First iteration of K-means
3. Creating file for selected batsman
4. Toss winning Vs match winning probability
5. Making program menu-driven.

**Lakshay Sharma**

1. Read data from csv and storing into array of structures.
2. First iteration of K-means
3. Creation of header files for linking
4. Testing and debugging

**DivyanshChandna**

1. Making Random Function for selecting initial random values.
2. Second and third iteration of K-means
3. Toss winning Vs match winning probability
4. Testing and debugging