

CSE 564 Project Proposal: CricVis

Pratik Panjwani (112681850) and Rajat R Hande (112684167)

Background & Problem Description

With an audience of over 2.5 billion, spread across 180 countries, **cricket** is the second-most **popular** sport in the world after football with every team having numerous players. But there is no holistic way of ranking the players based on their overall performance during the matches. The current available ranking models are based on either batting or bowling and off-late players are being ranked on both criteria(all-rounder). But given the amount of data available for cricket across formats spread across 3 different formats, there is ample scope for using this data for analyses. If we can develop a way of rating the players based on their past performances using the similar logic that is being used in today's cricket fantasy leagues that would help in selecting players for their fantasy league teams. Also, it is always difficult to compare players across eras given that the game, rules and formats change very quickly, so this can help in doing that comparison.

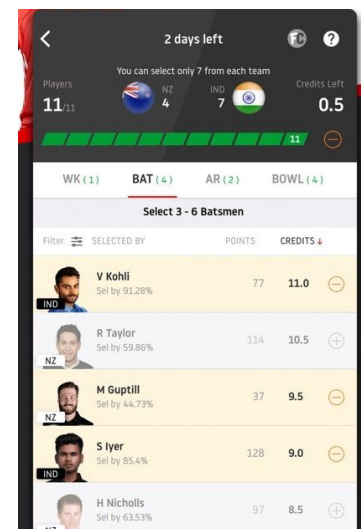
Objectives

We propose to develop a holistic approach for ranking these players using a point model inspired from famous fantasy leagues, where players are given ratings for runs scored, wickets taken, economy rate, fielding(catches and run-outs). We would be using historical data for cricket players and run our analysis on top of the data to generate the ratings for each individual player participating in the match. We would be creating visualizations in the form of a dashboard that would show the statistics for the player. The visualizations would help in better understanding the strengths of the players. A lot of insightful and interesting observations can be made using the cricket data available to us. Examples and use cases of such analytics are listed below:

- Batting Statistics Visualizations - Runs Scored per year, average, strike rate, boundaries, fifties and hundreds
- Bowling Statistics Visualizations - Wickets taken per year, average, strike rate, etc.
- Fielding statistics
- Similar form of rating players as done on some online fantasy gaming platform for cricket where users create a virtual team of real-life players and earn points based on the performances of these players in real matches.



A snippet with analysis during the world cup.



A popular cricket fantasy gaming platform

In our project we aim to use the below mens cricket dataset to create a visually appealing and informative dashboard that will provide the right facts and insights to the users and assist the users of platforms such as Dream11, CricketFantasy to make informed decisions.

Dataset Description

The data is available to be used from the following sources:

Dataset1: Data scraped from the API([link](#)) is available in json format. The data is available post 2011.

Dataset2: Cricinfo Statsguru Data is a well known site where all cricket related stats are available. A subset of this statistics are available on kaggle ([link](#)). Here the stats for batting, bowling and fielding are available in different CSV files.

In our Project, we would be joining the datasets for different activities(batting, bowling and fielding) across the 2 datasets mentioned above.

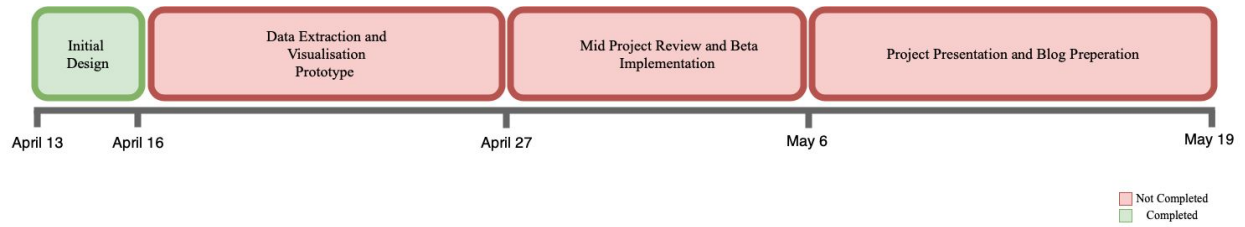
Some of the features we plan to use are:

- The runs scored across the span of the career (Indicating the current form).
- The strike rate and average runs scored/wickets taken by the batsman and bowler respectively.
- The number of boundaries hit.
- The fall of wicket of a particular player in different stages of the game.
- Mode of Dismissal of the batsman.
- The number of catches taken or run-outs affected by a player
- The fixture information of the teams and the player information for those matches.
- The venue and kind of pitch used for the game.

The following data records have been collected:

DataSet	Rows	Columns
Scoreboard	117694	33
Fixtures	19807	81
Players	18783	8
Teams	906	7
Venues	1006	7

Project Timeline



Prototype



References:

- [1] ESPN Crickinfo - <https://stats.espncricinfo.com/ci/engine/stats/index.html>
- [2] Kaggle-<https://www.kaggle.com/cclayford/cricinfo-statsguru-data#Men%20ODI%20Player%20Innings%20Stats%20-%2020th%20Century.csv>
- [3] Sample Dashboard - <https://knoema.com/wloaewd/odi-player-performance-card>
- [4] Fantasy League app - <https://www.dream11.com/>
- [5] Flask - <https://flask.palletsprojects.com/en/1.1.x/>
- [6] D3.js - <https://d3js.org/>