

Welcome to this hackathon on the Pro Kabaddi League. You may be wondering why we chose kabaddi as the theme for the hackathon and not our tried and tested football and cricket. Well, we at upGrad believe that learning and problem solving go hand in hand. In your data science career, you may come across situations where you are unaware of the background of a business but you still need to perform analysis on the data pertaining to that business. In such a situation, you will need to understand or gather all the information about the business and then proceed with your analysis. This is applicable to any field.

Feeling motivated yet?

No? Then listen to the [title track](#) for the Pro Kabaddi League. I am sure this will raise your enthusiasm.

Don't worry, we will first help you understand the game of kabaddi. Then, we will explain what needs to be done in this hackathon and, more important, how to win it. But remember, hackathons should be considered as more of a learning opportunity than a winning opportunity.

Let's get started with some basics of kabaddi. Here are some links that you can visit in order to familiarise yourself with the sport. We recommend that you go through the links in the same order that they are provided:

1. [Basics of Kabaddi](#)
2. [How to play Kabaddi.](#)
3. Video tutorial: [Link-1](#) and [Link-2](#)

Now that you have gathered all the information required to complete this hackathon, let's move on to the problem statement.

Interesting Facts about Kabaddi and the Pro Kabaddi League

- There have been three kabaddi world cups so far (the fourth one will be held this year), and India won the title all three times in the men's category. Interestingly, Iran was the runner-up in all the three world cups.
- The Indian women's kabaddi team also won all the three world cups organised so far.
- U Mumba is owned by Ronnie Screwvala. He is also the co-founder of upGrad Education.

Introduction

[Pro Kabaddi League](#) is a professional-level kabaddi league that started in 2014. Currently, the League is in its [7th season](#), which started on 20 July 2019 with the first match between U Mumba and Telugu Titans. The final match will be played on 19 October 2019.

In this hackathon, your task is to use your coding and analytics skills to predict various outcomes at the end of the tournament.

Tournament Details

Official Website	https://www.prokabaddi.com/
Start date	20-07-2019
End date	19-10-2019
Number of teams	12
Total league matches (including final)	137

Data

We will not provide you with a specific data set. You are free to use any data set as long as it is related to the Pro Kabaddi League. Remember, you will need to make predictions only for the current season of the tournament, i.e., season 7, but you are free to use the data from all the seasons.

You can find the data for all the seasons [here](#). You can refer to the links below to view team- or player-specific data.

Data Type	Source
All team and player data	https://www.prokabaddi.com/stats
Individual team data	https://www.prokabaddi.com/teams/bengal-warriors-profile-4 (On the official website, you can view the data of specific teams by clicking on the Teams tab.)
Individual player data	https://www.prokabaddi.com/players/maninder-singh-profile-143 (Once you choose a team, you can view all the data specific to a player by clicking on the player's profile and then selecting Stats.)
Alternative source	https://www.sportskeeda.com/go/pro-kabaddi/

How to load data into your system for analysis?

Well, this is not straightforward; in fact, it is one of the challenges of this hackathon. We have provided you with the links to get all the necessary data for your analysis. Now the challenge for you is to decide how to load the data into your system or software.

Here are some helpful links that you can use to tackle this challenge:

1. [For scraping the data using R.](#)
2. [Some analysis around the data.](#)
3. [For scraping using Python.](#)
4. [Using the jsoup library you can fetch necessary data by scraping web pages](#)

Apart from R and Python, you can use JavaScript, Java, or any other language or tool of your choice to scrape the data.

Tasks

So far, you have gathered information about the game of kabaddi and the Pro Kabaddi League, and also have access to various data sources that can be used for the analysis. The data set generated here could be extremely large. In that case, you can use the tools from the Big Data stack to perform your analysis. Now that you have the required data and also know how to handle a huge data set, let us help you understand the tasks involved in this hackathon and what you can do to win it. Let's RAID this!

Team Analysis

First, you will need to make predictions for the overall team by analysing the team data. Here are some team-specific predictions that you will need to make.

Task 1: Predict the winner of the tournament. (Max Points: 05)

Use the available data and predict which team will win season 7.

Criteria	Points Awarded
Prediction=Actual	05
Prediction=2nd Actual	03
Prediction=3rd Actual	01
For any other case	0

Here is the explanation for the above table. If you predicted that U Mumba would win season 7 and if your prediction came true, then you would be awarded 5 points. Now, if U Mumba came 2nd, then you would be awarded 3 points, and so on.

Task 2: Predict the top team in the points table after the completion of the league matches. (Max Points: 05)

Here, you will need to predict which team will top the points table after the league matches are completed (excluding the elimination rounds, the semi-finals and the final).

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

Here is the explanation of the above table. If you predicted that U Mumba would top the points table, and if your prediction came true, then you would be awarded 5 points; else, it will be evaluated for the 2nd position holder and 3rd. Easy, right?

You can access the current points table for season 7 here:

<https://www.sportskeeda.com/go/pro-kabaddi/points-table>

You can also use the following data for your analysis:

<https://www.sportskeeda.com/kabaddi/past-seasons-pro-kabaddi-points-table>

Task 3: Predict the team with the highest points for successful raids. (Max Points: 05)

A raid is considered successful when a player returns safely to his/her zone after tagging one or more players from the opposite team. Here, you will need to predict which team will earn the maximum points for successful raids.

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

We hope that by now you have understood the table fairly well. Moving on to the next tasks. hope that by now you would have understood the table fairly well

Task 4: Predict the team with the highest points for successful tackles. (Max Points: 05)

A tackle is considered successful when a raider is unable to return to his/her zone after tagging an opponent. Here, you will need to predict which team will earn the maximum points for successful tackles.

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

Task 5: Predict the team with the highest super-performance total. (Max Points: 05)

The super performance total (S.P.T.) of a team is expressed as:

S.P.T. = Total number of super-raids in the tournament + total number of super-tackles in the tournament + total number of all-outs inflicted in the tournament - total number of all-outs conceded in the tournament

You will need to predict the team with the highest S.P.T. in the tournament.

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

Player Analysis

Here, you will need to make predictions for individual players by analysing player-specific data. Here are the two predictions that you will need to make.

Task 6: Predict the player with the highest SUCCESSFUL RAID percentage. (Max Points: 05)

The successful raid percentage for a player is expressed as $(\text{Successful raids}/\text{Total raids}) \times 100$. You will need to predict which player will have the highest Successful raid percentage in the tournament.

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

Task 7: Predict the player with the highest SUCCESSFUL TACKLE percentage. (Max Points: 05)

The Successful tackle percentage for a player is expressed as $(\text{Successful tackles}/\text{Total tackles}) \times 100$. Here, you will need to predict which player will have the highest successful tackle percentage in the tournament.

Criteria	Points Awarded
Prediction = Actual	05
Prediction = 2nd position	03
Prediction=3rd position actual	01
For any other cases	0

Points Summary

Here is how the points are distributed among the different tasks in the challenge.

Task	Marks
Predict the winner of the tournament	05
Predict the the top team in the points table after the completion of league matches	05

Predict the team with the highest points for successful raids	05
Predict the team with the highest points for successful tackles	05
Predict the team with the highest super-performance total	05
Predict the player with the highest SUCCESSFUL RAID percentage	05
Predict the player with the highest SUCCESSFUL TACKLE percentage	05
Total	35

Submission

Here is a list of important points for all teams to consider before submitting their entries for the challenge:

- The teams should upload the data file and the code, along with the project report (README.md), to GitHub.
- The report should mention the outcomes predicted and the steps followed for making the predictions, starting with data selection. It should describe how the data was cleaned; how visualisations were done; how the modelling was approached; and how the results were evaluated.
- All your predictions should be backed with a detailed analysis and the code you used to make those predictions.
- Once the project is uploaded to GitHub, you will need to provide the details of the GitHub repository before the submission deadline. You should not change the GitHub repository after the competition deadline. Any changes found after the deadline would lead to immediate disqualification of the respective team(s), and they will no longer be eligible to participate in the challenge.

TERMS AND CONDITIONS:

- Challenge start date: 12th September 2019
- Registration for the challenge will close on 11th September 2019
- Challenge end date: 30th September 2019, 11:59:00 PM IST.
- No submissions/changes will be accepted after this date.
- Participating teams can have a strength of maximum 4 members and minimum 3 members, of which at least one is non-upGrad learners and should not have more than two upGrad learners.
- upGrad will not be responsible for the contribution of the team members or lack thereof.
- The solution for the challenge must be uploaded on the Github profile (here is a tutorial on [how to use Github](#)) of all the participants. The link for one of the uploads should be shared in the form posted on the learning platform.
- upGrad will only accept one submission per team.
- An expert panel formed by upGrad will take a decision on the winners. Decision taken by the expert panel will be considered final.

Eligibility Criteria

You must either be an upGrad learner or be directly invited by an upGrad learner to participate in this challenge.

Prizes

1st Prize: Amazon voucher worth Rs 16000 (per team) + Certificate from upGrad

2nd Prize: Amazon voucher worth Rs 12000 (per team) + Certificate from upGrad

3rd Prize: Amazon voucher worth Rs8000 (per team) + Certificate from upGrad

The prize will be divided equally between the members of the winning teams and sent by upGrad via email to each member of the winning team, irrespective of their contribution to the submissions.

In the event of a tie, the prize money will be equally distributed between the winning teams.



Other Terms and Conditions

All participants will be subject to and consent to the Terms & Conditions and Privacy Policy available on upGrad's platform at www.upgrad.com

You agree that your submissions, in whole or in part, do not infringe any third party intellectual property rights. If you are found to be in breach of the same, you and your team members will be severely and jointly liable for the same.

You are solely responsible for any and all expenses that you incur in entering the competition and you will not be reimbursed regardless of whether or not you win the challenge.

You grant upGrad a worldwide, perpetual, royalty-free, irrevocable license to use the submissions for any purpose it deems fit, including for publicity and advertisements.

upGrad has the exclusive right, at its sole discretion, to disqualify any participant or team, if they are found to breach these Terms & Conditions.

