

# **IPL contest - FAQs**

# Files available on the website

Registration is now closed.

Available at this link [https://internalapp.nptel.ac.in/contest/contest\\_details.html](https://internalapp.nptel.ac.in/contest/contest_details.html)

1. Data set - Data of 1600 innings of past T20matches
2. Input format - Format of input file
3. File structure - denotes the files to be zipped and submitted, contains main.py and predictor.py
4. Requirements.txt contains the libraries that have been added to the test environment and which alone can be used in the submitted code
5. Players list - has the list of ALL players playing in the current tournament
6. Sample input files - has a sample for every match that is to be played between the two teams

Inside your dashboard

Actual input file used for score prediction will be made available the day after the match is over.

# Information is publicly available - all details given on Contest details

The screenshot shows a web browser window with the address bar containing `onlinedegree.iitm.ac.in/contest/`. The browser's address bar includes navigation icons (back, forward, refresh) and a star icon for bookmarks. Below the address bar, there are icons for various applications: 'Apps', 'Gmail', 'Podcasts', 'nptel', and 'swayam'. On the right side of the browser, there is a 'Readin' icon and a user profile icon labeled 'B'. The website header features the IIT Madras BSc Degree logo on the left and a navigation menu with the following items: 'Home', 'Timeline', 'Forum', 'Leader Board', and 'Contest details' (which is highlighted with a black border). Below the navigation menu is a dark red banner with the text 'CRICKET AND CODING - CRACK THE CHALLENGE' and a white button labeled 'PARTICIPATE NOW'.



## Q. How will you give input file like separate files for first and second innings?

Yes, every innings will have a separate input file which we will create after the actual match is over. The format of the input file is given on the website for your reference.

venue	innings	batting_team	bowling_team	batsmen	bowlers
MA Chidambaram Stadium	1	Mumbai Indians	Kolkata Knight Rider	Quinton de Kock, Rohit Sharma, Suryakumar Yadav	Harbhajan Singh, Varun Chakravarthy, Shakib Al Hasan, Pat Cummins

**Q. Without knowing who are playing in the first 6 overs, how is it possible to predict the score?**

**The submission for competition is open now but will the input file be provided?**

We have shared a sample input file for every match played by the 2 teams. This can be used to test your code for the match of the day to give the prediction. We will not be using this score anywhere.

Only the score returned by your code run on the actual input file by us will be counted towards the contest.

We will share the actual input file with you the next day. You can run it against your code to see if you get the score we also have displayed, which will confirm the evaluation is correct.

## Q. Does the input have ball by ball data for the first 6 overs everyday or just batsmen and bowler info ?

The input file does not have ball by ball information, just the batsmen and the bowlers of the first 6 overs.



venue	innings	batting_team	bowling_team	batsmen	bowlers
MA Chidambaram Stadium	1	Mumbai Indians	Kolkata Knight Rider	Quinton de Kock, Rohit Sharma, Suryakumar Yadav	Harbhajan Singh, Varun Chakravarthy, Shakib Al Hasan, Pat Cummins

### Kindly note:

1. The sample input file is not a (real) match data.
2. If the player is new to the tournament, then the player name will be similar to the official website.
3. The original input file for each match will be updated after the end of the match.

**Q. In inputFile.csv the batsmen name are given as Rohit Sharma but in training data it is given RG Sharma. Similarly, for all batsmen and bowlers.**

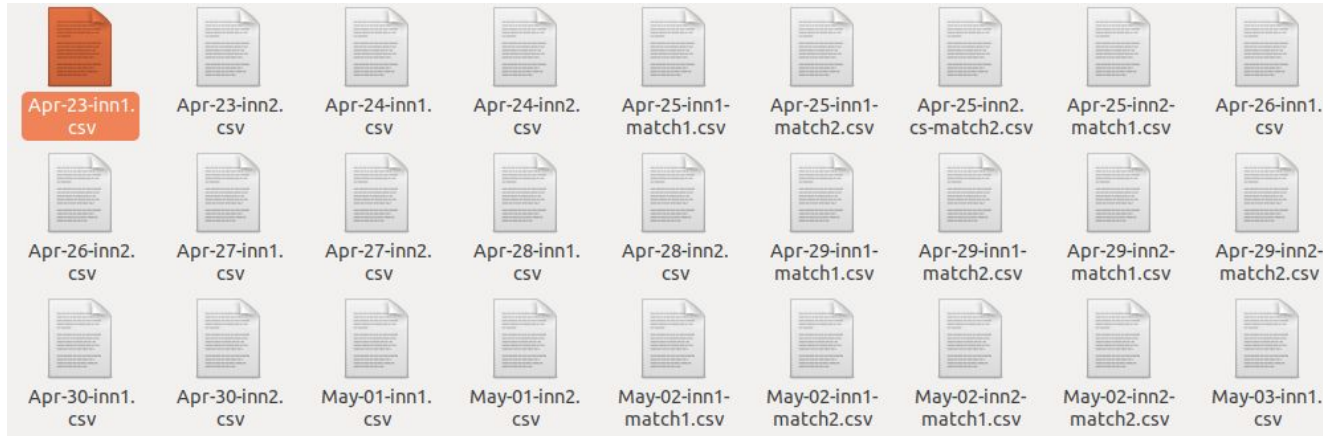


Players
AM Rahane
RA Tripathi
MJ McClenaghan
SPD Smith
SL Malinga
MK Tiwary
JJ Bumrah
KV Sharma
KH Pandya
HH Pandya
MS Dhoni
LMP Simmons
PA Patel
JD Unadkat
Washington Sundar
LH Ferguson
SN Thakur

Just ensure that the current player names are matched against the names given in the data set.

## 5. Whether it is allowed to create our own input file ( with single input data) to testing my model?

Yes. You can create sample/dummy input files to test your model based on the past data. Ensure the format we have given is followed so that code can be run by us.





**Q. While submitting the code(zip)/algorithm, there are mandatory fields of Predicted scores. Will you run the code yourself post match and consider that output or will it be the one submitted in the google form pre match.**

These fields have been provided for you just to make sure that your code works. These scores will not be used for the evaluation. - We are in fact removing it from tomorrow's form.

Only the score returned by your code run on the actual input file by us will be counted towards the contest.

## Q. Where do I find the ball by ball dataset?

[↓ DATASET](#)[↓ INPUT FORMAT](#)[↓ FILE STRUCTURE](#)[↓ REQUIREMENTS.TXT](#)[↓ PLAYERS LIST](#)[↓ SAMPLE INPUT FILES](#)[335982.csv](#)[335983.csv](#)[335984.csv](#)[335985.csv](#)[335986.csv](#)[335987.csv](#)[335988.csv](#)[335989.csv](#)[335990.csv](#)[335991.csv](#)[335992.csv](#)[335993.csv](#)[335994.csv](#)[335995.csv](#)[335996.csv](#)[335997.csv](#)[335998.csv](#)[335999.csv](#)

match_id	season	start_date	venue	innings	ball	batting_team	bowling_team	striker	non_striker	bowler	runs_off_bat	ex
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.1	Kolkata Knight Riders	Royal Challengers Bangalore	SC Ganguly	BB McCullum	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.2	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.3	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.4	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.5	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.6	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	0.7	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	P Kumar	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	1.1	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	Z Khan	0	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	1.2	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	Z Khan	4	
335982	2007/08	2008-04-18	M Chinnaswamy Stadium	1	1.3	Kolkata Knight Riders	Royal Challengers Bangalore	BB McCullum	SC Ganguly	Z Khan	4	

**Q. We have submitted a code. How to know, if it works fine without errors?**

Everyday we update the result of the code execution on your dashboard. If the code worked, we update the score predicted. Else we display the error obtained.

**Q. So, how exactly will our code be tested? Will you add the right inputFile (after the match) into our folder and then run main.py file because the code assumes the presence of the inputFile in the same folder.**

Yes. We will place the actual input file at the right place in the folder and will run your code.

**Q. The team members of my team are not able to log into the contest page.**

For every team - there is only one email id that is registered and all members need to login only through that. Each team member does not have different login credentials.

**Q. I've seen the requirements.txt file that says the packages I'm supposed to be using. There are multiple commonly used regression packages such as XGBoost, LightGBM which are not mentioned over there. Is it okay if I use them. Even 'joblib' which is used in the tutorial video is not mentioned in the file.**

Only the packages given in the requirements.txt file have been loaded into the test environment. Any other packages you use will not be available for use. Default packages (which do not need installation) like joblib, pickle may be used.

## **Q. How to know if our submission has been accepted**

Your scores will be made available in your dashboard the next day by 5 PM (we will try to publish as soon as we can).

If there are any errors, they will be displayed.

**Q. Do we need to submit the file daily although we did not make any changes?**

You can make changes in your code to see how to reduce the difference between the predicted and actual score and resubmit it. Your latest submission will be considered for evaluation. So you can update your program regularly but it's not compulsory.

If you submit only once, that submission will be run against the input file every day.



**Q. I have not received any confirmation mail after registration. How would I know if I have registered successfully for the competition?**

Please login with your Google enabled email id (email id with which you registered), and if you can see your dashboard, then you can be sure that you have successfully registered.

**Q. Is it necessary to include inputFile.csv in the zip folder? If yes, do we have to modify the inputFile. csv ourselves before every submission?**

This is kept in the folder so that errors pertaining to path do not arise. You may modify it to test your code for the day.

We will be using the actual input file containing the maths details for the corresponding day.

**Q. There were certain discrepancies in the venue names which I had to clean and replace with correct names but some venue names were quite similar like M Chinnaswamy Stadium and M. Chinnaswamy Stadium. The difference is just of a dot after M but this would throw an error. What will be done in such a case?**

Kindly refer to the 'sample input files' for getting the standard venue names available here:



You may accordingly preprocess the dataset.

**Q. The venue format provided in the input file will be "name of the stadium" or "name of the stadium, place"?**

It will be the name of the stadium only.

**Q. Should we give the input file according to the feature engineering we did or do the IITM gives the input file?**

The format of the input file will be as shown on the contests details page. You may perform feature engineering in your code.

**Q. Does the PredictRuns() function have to return an int value for it to be considered correct? My linear regression model outputs floating point values that I round before returning the value.**

It should be int. That is what the expectation is.

**Q. Since we need to use 2 input file per match for the 2 batting team. So does that mean we need to upload 2 zip files each with predictor, main and one of the input files?**

You only need to upload one zip file (containing any one of the sample input files)

**Q. Can I save my model as .pickle or .joblib file when training the model using scikit-learn package.**

Yes, you may use these two packages.

**Q. Can we submit our code files as .ipynb notebooks?**

**Q. I have made the entire code on kaggle, for some reason I am not able to convert it to the required submission format, is there any way I can submit it as it is?**

No, you need to submit the code in the required format (as shared on the contest details page).

**Q. Will the input file be the excel sheet or only the entries?**

It will be a csv file as shared on the contest\_details page.

**Q. We registered for the contest on 25th April as a team of four members. We didn't get anything like some videos introducing Data Science or the platform and resources we have to use for code submitting. Neither we got the datasets for the challenge. It make us feeling like we have missed something. Could you please share this information?**

You need to access these details from the context\_details page. Videos and dataset etc are available there:

[https://internalapp.nptel.ac.in/contest/contest\\_details.html](https://internalapp.nptel.ac.in/contest/contest_details.html)



**Q. I had queries regarding the contest whether code has to be submitted daily or on 26th May and if daily what is the deadline time for it?**

You may keep updating your code if you wish to improve it based on the scores that you obtain.

The deadline time for daily upload is 2 PM on days when there are 2 matches.

If there is only one match, the deadline is 7pm.

**Q. While reading csv file we have given path when you are trying to run our code in your system it will generate error because of path. So how to overcome that problem?**

You may use relative paths rather than absolute paths to avoid path errors.