₹

?

✔

🏆

(/contest/click-prediction/lb)

Registered

Starts at	Fri Sep 15 2017 00:00:00 GMT+0530 (IST)
Closes on	Sun Oct 15 2017 23:59:59 GMT+0530 (IST)
Mode	Online
Fee	Free
# Participants	2975
Prizes	Prizes worth 2 lakh INR and top 20 will get discounts on DataHack Summit tickets

Click Prediction Hackathon

Hackathons are becoming a popular way for individuals to showcase their machine learning expertise and knowledge. Hackathon mania is a new fever amongst data scientists and machine learning enthusiasts.

Come be part of this competition that you will give you more reasons to love hackathons.

Prizes:

Top 3 rankers of the hackathon will be awarded

- Rank 1: INR 1.25L
- Rank 2: INR 50K
- Rank 3: INR 25K

This is not all, if you get in the top 20, you will get exclusive benefits on India's Largest Data Science Event, in Bengaluru, DataHack Summit 2017 (https://www.analyticsvidhya.com/datahacksummit/)

- Top 10 contestants will get a 25% OFF on Conference Tickets (https://www.analyticsvidhya.com/datahacksummit/all-tickets/) (2 days and 3 days Pass)
- Contestants from ranks 11-20 will get a 15% OFF on Conference Tickets (https://www.analyticsvidhya.com/datahacksummit/all-tickets/) (2 days and 3 days Pass)

Rules

Contest Guidelines

- One person cannot participate with more than one user accounts.
- Appropriate taxes will be applicable on the prize money.
- You cannot use ID as a feature for modelling
- You cannot use future data for modelling
- You should submit a .zip for submission

Tools

- You are free to use any tool and machine you have rightful access to.
- You can use any programming language or statistical software.

Solution Checker

- You are free to use solution checker as many times as you want.
- Adding comment is mandatory for use of solution checker
- Comments will help you to refer to a particular solution at a later point in time.

Final Submissions

- Setting final submission is mandatory. If you don't make final submission, your entry would be dis-qualified.
- No submissions would be entertained after the hackathon ends.
- Code file is mandatory while setting final submission. For GUI based tools, upload zip file of snapshots of steps taken by you, else upload code file.
- The code file uploaded should be pertaining to your final submission. If we find any discrepancy between the two, your entry would be dis-qualified.

Team formation

- Maximum of 2 people can form a team.
- In case a team wins, prize would be distributed equally among team members
- Team once created can't be dissolved.
- Teams can't be merged.

Expected conduct

- At any point in the hackathon, you are expected to respect fellow hackers and act with high integrity.
- Slack Live Chat admins hold the right to blacklist / block any participant found to use foul / disrespectful language. Chat forum will be closely monitored.
- Analytics Vidhya holds the right to disqualify any participant at any stage of competition if found indulged in fraudulent practices.

Registration Fee

Free

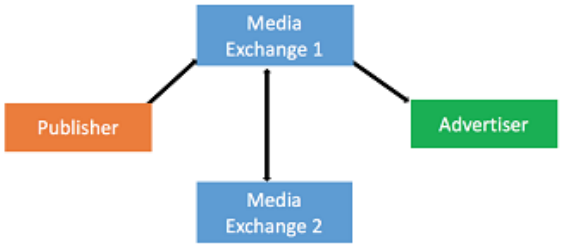
Problem Statement

Welcome to Collectcent Click Prediction Challenge!

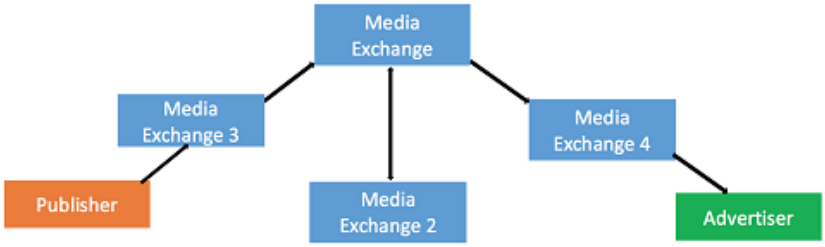
At Collececent, Technology is at the heart of what we do. We believe in providing the best experience and returns for our publishers & advertisers for their investments and partnerships with us.

We deliver more than 3 Billion clicks per month to our advertisers delivering 4.5 Million monthly sales events. We want to leverage machine learning to improve the conversions for our customers.

This is how an Advertisement network works in general:



The network buys space on the Publishers site and then shows an advertisement about the Advertiser at that space. The advertiser pays the network for every conversion from the clicks. The network in turns pays to the publisher after keeping it's commission.
Note: In real life, publisher and advertiser can also act as a media exchange



We want to open this challenge to Analytics Vidhya community to bring out the best models predicting conversions on our clicks. More specifically, you have to predict the values of "ConversionPayOut". Needless to say - the person with the best model wins.

All the best!

Evaluation Metric

The evaluation metric for this competition is weighted RMSE

Data

The sample data is provided of 14 days, from 21st August to 3rd September. The train file (1.1GB size) contains sample data from first 10 days whereas you are asked to predict on sample data of last four days in the test file (~0.5GB). Also, the public leaderboard would take into account the first day of test predictions, whereas private leaderboard will take into account rest three days of test predictions. Below is the data dictionary:

Variable	Description
ID	Unique ID of click
Country	Country Code
Carrier	Wireless Network Operator Code

TrafficType	Whether the advertisement is for Adults or Mainstream
ClickDate	Date at which the advertisement was clicked
Device	Type of Device from which advertisement was clicked
Browser	Type of Browser from which advertisement was clicked
OS	Type of OS from which advertisement was clicked
RefererUrl	Url of Source website
UserIp	IP of User who clicked
publisherId	Unique ID of publisher
subPublisherId	Unique ID of sub publisher
advertiserCampaignId	Unique ID of campaign of advertisement
Fraud	If the click was fraud or not
ConversionStatus	Was the click Converted or not
ConversionDate	Date of Conversion if Conversion happens
ConversionPayOut	Pay Out of Conversion if Conversion happens (in rupees)

Test File (/contest/click-prediction/download/test-file)
Train File (/contest/click-prediction/download/train-file)
Sample Submissions (/contest/click-prediction/download/sample-submission)

Solution Checker

This contest is over.

Join Slack Live Chat

(https://analyticsvidhya.com)