Problem Statement | Discussions | Our Submissions | About Hackathon

Conversion fraud in Digital Advertising

Digital Advertising is changing at a rapid pace with a huge increase in digital audience. At the same time, the digital advertising success metric is shifting from audience volume (eg. Impression count) to conversions (eg. lead submissions) as the success metric. This requires higher transparency and control on the conversions.

Colombia, the digital advertising arm of Times Internet Limited has seen significant growth in its digital advertising inventory. It wants to ensure that in all its conversion-based campaigns, no unfair advantage is given to the publishers generating fake leads.

Your task is to segregate the test data between genuine and false conversions by identifying the maximum possible leads generated by the malignant technique.

Note-
Joining with Click log:
imprId (Click Log) and imprid_cr (Conversion Log - Test and Train Data)

(Use click log data for additional data required for identifying conversion fraud)

Essential Columns

- client id: Advertiser ID
- pubclient id: Publisher ID
- clickIp: IP Address
- clmbuser id : unique user id
- impr id: Unique Key for every served impression
- site id: Publisher wesite
- goal id: Conversion`s goal type identification id
- City id / State id / CountryDim id: Geo Details
- browser id: browser used for accessing publisher on any device on web.
- adslot id: slot id where advertisement is displayed on any site (unqiue for all sites)
- crtd: timestamp of the action
- itmclmb id: Image/Creative shown
- ispDimId: Internet Service Provider
- devTypeDimId: Device Id
- osVerDimId: OS Version

Data Set Download Data Set

File Name	Description	Format	Size
click_log.csv	Click logs	csv	
Training Data.csv	Train data	csv	
sample_submission.csv	Sample submission	csv	
Test Data.csv	Test data	csv	

Data Dictionary

Here's a brief version of what you'll find in the data description file.

Variable	Description
record_id	Unique record id
conversion_fraud	conversion_fraud column as True for marking a conversion entry as fraud

Submission

1. Submission file
2. Presentation (Download sample ppt)

Evaluation Metric

Your score will be calculated based on Accuracy.

Upload Source Files (Optional)

You need to submit a zip or tar archive consisting of a text file explaining your approach, details about feature engineering, tools you used and the relevant source files.

Upload File

Upload Submisssion

Your submission should be in CSV format. You can upload this in zip archive, if you prefer.

Drag and drop or upload your submission file below

Upload File