Data Desription :

File descriptions

* train - Training set. 10 days of click-through data, ordered chronologically. Non-clicks and clicks are subsampled according to different strategies.
* test - Test set. 1 day of ads to for testing your model predictions.
* sampleSubmission.csv - Sample submission file in the correct format, corresponds to the All-0.5 Benchmark.

Data fields

* id: ad identifier
* click: 0/1 for non-click/click
* hour: format is YYMMDDHH, so 14091123 means 23:00 on Sept. 11, 2014 UTC.
* C1 -- anonymized categorical variable
* banner\_pos
* site\_id
* site\_domain
* site\_category
* app\_id
* app\_domain
* app\_category
* device\_id
* device\_ip
* device\_model
* device\_type
* device\_conn\_type
* C14-C21 -- anonymized categorical variables