

Kaggle Project

Click- Through Rate Prediction

<https://www.kaggle.com/c/avazu-ctr-prediction>

The detailed description of this project is available on the website link provided above.

This problem has huge train and test datasets which might not be feasible to process on personal laptops and hence we have taken 3 parts of the original train dataset to provide you one part as train data to build your models on, one part as test data for you to test and evaluate your models on and one part will be used by us to test you models on when you submit the models (Note: the 3rd part will not be provided to you).

Since you will not be using the original test data provided on the Kaggle website, you cannot submit this project to Kaggle and hence are expected to save your models in your Microsoft Azure ML account. We would login to your account and run the model on the test data (the 3rd part mentioned in the above paragraph).

The Click- Through Rate Prediction could be considered as a Classification problem. Microsoft Azure Machine Learning Studio has good two-class classification algorithms which could be used for this problem. You can use its tools to do some data preprocessing. You have simple training and testing sets, you can have a look at them and decide what approach you need to follow. No method is wrong. You can get a solution using any method you think is right. You are expected to explore and experiment with the various algorithms available in the tool to decide which one or ones suit your data the best.

You will be evaluated based on your approach and your efforts taken to get a good score. You need to adhere to the phases of the project submission to get full credit. You will also have to present the project in class which contributes to your grades. You also need to write a report about your project.

The phases of the project and the format of the report are uploaded to the Blackboard.

Best of luck!