

## Goal

The goal of this exercise is to predict conversion rate by device for a set of entities.

## Definitions

Conversion rate = the probability of a conversion event given a click event

## Data

The data can be found here:

[technical\\_test\\_training\\_data](#) It is a CSV file compressed using xz compression.

The data contains the following columns

- entity\_id = the ID of the entity
- att1160 (160 columns) = attributes/features of the entities
- device = the device that the clicks/conversions occurred on Computer, Smartphone or Tablet
- Clicks = the number of clicks
- Conversions = the number of conversions

## Required output

A CSV file containing the following columns

- entity\_id
- device
- prediction

entity_id	device	prediction
1	Computer	0.05
1	Smartphone	0.05
1	Tablet	0.05
2	Computer	0.05
2	Smartphone	0.05
2	Tablet	0.05

Additionally, please provide the code you used to solve the problem.

## Evaluation of the results

The results will be evaluated by comparing the predictions to a test data set. To succeed, the mean squared error (MSE) of predictions vs. observed data will have to be lower than that of the following naive methods

1. predicting 0 for all entities
2. predicting the mean of the entire train data set
3. predicting the mean of the train data set by device
4. predicting the mean of the train data set by entity\_id