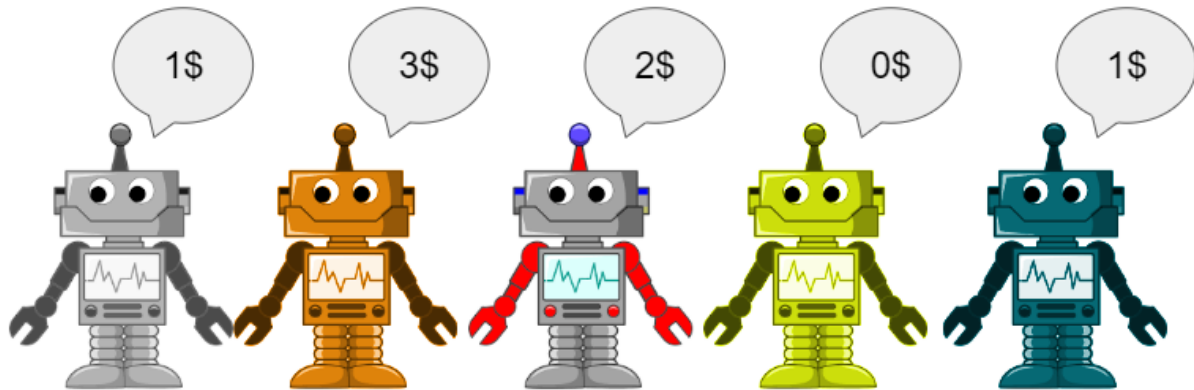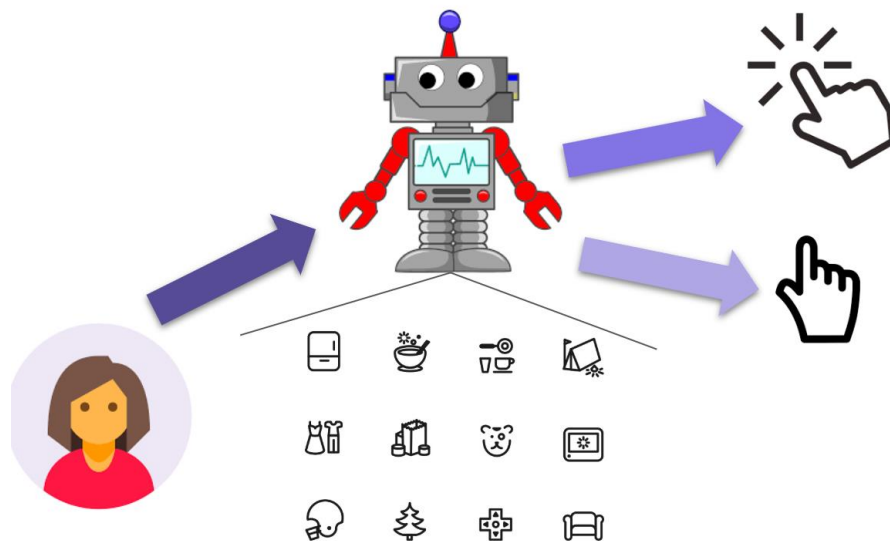# Click Prediction Challenge

Understanding the CPC (cost per click) market place

Every time a user opens a mobile app, an auction starts between several agents.
The winner of the auction gets to display an ad to the user.

However, the agent gets paid only if the user clicks on the ad

In order to make an informed decision, each agent needs to take the users features into account and make a prediction.

The Dataset

The data is available for download at this link:
https://drive.google.com/open?id=1fSuT4hR-HHmXi4NILivrrdXSU5rAm4lE

The data is given in CSV format, with the last column being our target variable (click / skip).

The data is taken from an ad marketplace, between September 2018 and November 2018, and the data is filtered on bid-requests of Android users from the United states.

Each row is a bid-request has the following user features:

App id	User state	User isp	Device maker	Device model	Device osv	Device height	Device width

As well as 2 special fields:
- **Utc_time** - the number of mili seconds that passed since January 1st 1970
  - A common time format, also known as "epoch"
- **Click** - Our target variable
  - 1 - if the user clicked
  - 0 - if the user ignored our ad

Task

You are required to design a bidding strategy given the dataset at hand.

We would like you to follow these steps:
1. Rank existing features importance
2. Engineer new features (e.g. day-of-week)
3. Optional: Enrich the dataset with external sources (e.g. timezones, app categories)
4. Predict whether the user clicked the ads or not (evaluate the performance) using the features from questions 1, 2, 3.

Tips:
1. If you split the data to train and test sets, do it wisely and not at random
2. A combination of features could be more helpful than a single feature prediction.
3. If you filtered data out, make sure you have enough records to have a statistical significant prediction.

Submission

You may use any programming language you'd like, as long as your code is well documented. We highly recommend to use jupyter notebook for your exploration process, this will make life easier to bind the code, plots, and results in a single notebook.