

# Industrial project description "Click-through rate prediction"

Answer the question to outline your project. Choose one of the roles: an Expert or an **Analyst**.

Chosen role: Analyst.

## 1 Planning the industrial research project

Before planning the research, the analyst and (**expert**) discuss the key issues. After the long dash — our remarks.

1. Goal of the project. (**Expected development result.**) — The main goal of the project is to examine the relationship between different advertisements and clicks on them and also to optimize online advertisement campaigns. We should answer: what is the best decision about ad placement and design?
2. Applied problem solved in the project. (**How will the result be used?**) — Construction of a model to predict the likelihood that an ad will be clicked in order to maximize their return on investment and improve user experience.
3. Description of historical measured data. (**Formats and timing.**) — The data for this project is a table with information about ads, their main properties and clicks. The rows of the table are ads and the columns are properties: clicked or not, date of a publication, description of the ad, description of a site or application where the ad was placed, information about users, and other useful features.
4. Quality criteria. (**How is the quality of the obtained result measured, what is in the report?**) — As analysts, we must maximize likelihood and accuracy of prediction. For business, the main quality criteria is the return on investment (ROI) of the ad.
5. Project feasibility. (**How to show that the project is feasible, list of possible risks.**) — For analyst, possible risks related to wrong hypotheses about data, false correlations which can lead to decreasing accuracy of model and unnecessary costs for business.

6. Conditions necessary for successful project implementation. (**Organization of work.**) — The conditions are as follows: a sufficiently large data set, computing power (CPUs, GPUs if necessary), competent communication between the contractor and the customer.
7. Solution methods. (**Procedure libraries.**) — Exploratory data analysis, visualization, different linear models for predicting, gradient boosting.

## 2 Research or development?

In other words, novelty or technological advancement?

Analyst: What impact will the research have on the field of knowledge? How useful will it be?

This research can be useful for developing new approaches and models to improve the quality of solutions. With ever-increasing amounts of data, the demand for more productive and sophisticated models that better extract patterns from data is also increasing. Moreover, the results of such a study can be used in related fields, for example, in the construction of recommendation systems.

Expert: (**How long will the model be used? What will replace it in the future?**)