# Project Evaluation

To complete the final sprint successfully, you'll need a score of five story points (SP). These are conventional units for measuring task difficulty. Get:

* from 4 to 6 SP for the main project;
* 1 SP for the additional task.

You'll build a prototype of a machine learning model. These will be your steps:

1. Make a work plan. When you first see the assignment, you'll notice that it's incomplete and contains unnecessary information. Do an exploratory analysis of data to find out what questions to ask.
2. Research the task. Ask your team leader questions.
3. Develop a model. Send your code to the project reviewer.
4. Prepare the report. Send the report to your team leader to check if you completed the tasks successfully.

The final score depends on the quality of your model. The community manager will share the evaluation criteria with you after step 2.

The additional task is based on the same data as the main project.

At the end of the sprint, the team leader will review your code.

## Metric

Target feature: The 'EndDate' column is equal to 'No'.

Primary metric: AUC-ROC.

Additional metric: accuracy.

Evaluation criteria:

* AUC-ROC < 0.75 — 0 SP
* 0.75 ≤ AUC-ROC < 0.81 — 4 SP
* 0.81 ≤ AUC-ROC < 0.85 — 4.5 SP
* 0.85 ≤ AUC-ROC < 0.87 — 5 SP
* 0.87 ≤ AUC-ROC < 0.88 — 5.5 SP
* AUC-ROC ≥ 0.88 — 6 SP

# Project Description

Telecom operator Interconnect would like to be able to forecast its customer churn rate. If a user is found to be planning to leave, they will be offered promo codes and special plan options. Interconnect's marketing team has collected some of its customers' personal data, including information about their plans and contracts.

## Interconnect Services

Interconnect primarily provides two types of services:

* Communication by landline. The phone can be connected to multiple lines simultaneously.
* Internet. The network can be configured via a telephone line (DSL, *digital subscriber line*) or via a fibre optic cable.

Some other services offered by the company include:

1. Internet security: Antivirus software (*DeviceProtection)* and a blocker for malicious websites (*OnlineSecurity*).
2. A technical support line (*SoporteTécnico).*
3. Cloud file storage and data backup (*BackupOnline*).
4. Streaming TV (*StreamingTV*) and Movie Directory (*StreamingMovies*)

Customers can choose between a monthly payment or sign a 1 or 2-year contract. You can use multiple payment methods and receive an electronic invoice after a transaction.

# Data Description (Data Set)

The data consists of files obtained from different sources:

* contract.csv — contract information;
* personal.csv — personal data of the customer;
* internet.csv — information on Internet services;
* phone.csv — information about telephone services.

In each file, the customerID column contains a unique code assigned to each customer. The information in the contract is valid as of February 1, 2020.



The data is also located on the platform, in the folder '/Datasets/Final\_provider/'.

# Task 1 – Additional Information

1. Make a work plan.
   1. Before you begin the full development of the project, you will need to perform an exploratory data analysis. For that, at the end of the Jupyter Notebook, he writes:
      1. A list of clarifying questions you want to discuss with your team leader.
      2. A rough plan for solving the task, including 3 to 5 key steps explained in one or two statements each.
   2. 🔎 Important: The team leader will review your questions and work plan. Questions will be answered during a video call. The code will be reviewed only if there are doubts.

# Task 2 – Research the Task

N/A. Included in task # 1.

# Task 3 – Preparation of the Model

Upload your solution code.

Here are some of the criteria that project reviewers use:

* Did you complete all the steps?
* Does the solution match the condition of the task?
* Did you train and test the model correctly?
* What is the quality score of the best model?
* Did you maintain the structure of the project?
* Did you keep the code clean?

# Task 4 – Report on the Results

At the end of the *Jupyter Notebook*, make a report with the solution. The team leader will review it. The code will be reviewed by the team leader only if there are some doubts.

In the report, he answers the following questions:

* Which steps of the plan were performed and which steps were omitted (explain why)?
* What difficulties did you encounter and how did you manage to solve them?
* What were some of the key steps in solving the task?
* What is your final model and what level of quality does it have?

Here are some of the criteria used by the team leader:

* Did you answer all the questions?
* Are the answers clear?