Project description

The telecom operator Interconnect would like to be able to forecast their churn of clients. If it's discovered that a user is planning to leave, they will be offered promotional codes and special plan options. Interconnect's marketing team has collected some of their clientele's personal data, including information about their plans and contracts.

Interconnect's services

Interconnect mainly provides two types of services:

- Landline communication. The telephone can be connected to several lines simultaneously.
- Internet. The network can be set up via a telephone line (DSL, *digital subscriber line*) or through a fiber optic cable.

Some other services the company provides include:

- Internet security: antivirus software (*DeviceProtection*) and a malicious website blocker (*OnlineSecurity*)
- A dedicated technical support line (*TechSupport*)
- Cloud file storage and data backup (*OnlineBackup*)
- TV streaming (*StreamingTV*) and a movie directory (*StreamingMovies*)

The clients can choose either a monthly payment or sign a 1- or 2-year contract. They can use various payment methods and receive an electronic invoice after a transaction.

Data Description

The data consists of files obtained from different sources:

- `contract.csv` — contract information
- `personal.csv` — the client's personal data
- `internet.csv` — information about Internet services
- `phone.csv` — information about telephone services

In each file, the column `customerID` contains a unique code assigned to each client.

The contract information is valid as of February 1, 2020.

Dataset

You can download the dataset by following this link.

The data is also available for usage on the platform in the `/datasets/final_provider/` folder.

Clarification: Summary

Target feature: the `'EndDate'` column equals `'No'`.

Primary metric: AUC-ROC.

Additional metric: Accuracy.

Assessment 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

Final Project: Work Plan

You're going to have to preform an exploratory data analysis. At the end of *Jupyter Notebook,* write:

A list of clarifying questions
A rough plan for solving the task, specifying 3-5 basic steps and explaining each step in one or two sentences

The expert tutor will check your questions and work plan. The code will be reviewed by the expert tutor only if there are some areas of doubt.

Final Project: Solution Code

Upload the code of your solution.

Here are some of the criteria used by the project reviewers:

Have you performed all the steps?
Does the solution match the task condition?
Have you trained and tested the model correctly?
What is the quality score of the best model?
Have you kept to the project structure?
Have you kept the code neat?

Final Project: Solution Report

Make a report at the end of the *Jupyter Notebook* with the solution. The expert tutor will check it. The code will be reviewed by the expert tutor only if there are some points of doubt.

In the report, please answer the following questions:

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

Here are some of the criteria used by the expert tutor:

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