Proposal for the candidate

Puzzle Overview

As a credit company, it is important to know beforehand who is able to pay their loans and who is not. The goal of this puzzle is to build a statistical/machine learning model to figure out which clients are able to honor their debt.

Structure

Besides these instructions, the ZIP file contains 2 datasets in CSV format:

- *puzzle_train_dataset.csv*: training dataset
- *puzzle_test_dataset.csv*: test dataset. It contains the same columns as the training dataset, except for the `default` variable

Your goal is to predict default, which is identified by the *default* variable in the training dataset.

Deliverable

You must send us an email with:

- an explanation of what you did (maximum length of 2000 characters with spaces)
- your own code, which we should be able to run (you can assume that we have access to the same resources that you used)
- instructions about how to run the code (e.g., software, programming language, required dependencies, commands to run)
- a *.csv* file called *predictions.csv* with 2 columns: the IDs from *puzzle_test_dataset.zip* and the predictions labeled by *predictions*.

Please try to follow these instructions as closely as you can.

What we will evaluate

Your ability to:

- handle/clean data
- write good quality code (e.g., reproducible, readable)
- apply machine learning models to real problems
- split complex and real problems into solvable pieces