

Information on the final project

Description:

This project counts for 20% of the final grade and is to be done in teams of no more than 2 or 3 people.

It consists of an analysis of a case study, to be chosen among three choices (see pdf files for case descriptions):

- Credit Risk Case Study (Credit_Risk.pdf)
- Fundraising Case Study (Fundraising.pdf)
- Predictive Maintenance Case Study (Predictive_Maintenance.pdf)

Due Date:

December 19 by 11:55 p.m. on Zone Cours

Project Deliverables:

- A report of up to 7 pages describing your analysis strategy, including data cleaning, exploration, pre-processing, modeling, and conclusions/interpretation of results. The report should not contain R code, but may contain figures and tables (only if they are relevant and provide useful information that you comment on). The report may also contain result metrics that you must interpret. The report can also contain information about strategies that you tried and thought were really relevant (explain why), but did not work.
- A results file depending on the project chosen:
 - Credit risk project: a csv file called `Result_credit.csv` that contains 2 columns. The first column called `TRAIN_ID` contains the IDs from the `CreditGame_Application.csv` file. The second column called `AMOUNT` contains the amount of credit granted (0 if the credit is not granted).
 - Fundraising project: a csv file called `Result_collection.csv` which contains 2 columns. The first column called `ID` contains the IDs of the `MembersList.csv` file. The second column called `CALL` contains the value 1 if the client should be called and 0 otherwise
 - Project on predictive maintenance: a csv file called `Result_maintenance.csv` which contains 2 columns. The first column called `ID` contains the ID of the `sensors_score.csv` file. The second column called `MAINTENANCE` contains the value 1 if the pump should be serviced at 6 months, and 0 otherwise.
- Your complete R code for the model and the final strategy chosen should be provided in a separate file. This code should be clean and should not contain any descriptive analyses you have done or any of your tests. It should contain only the steps necessary to answer the question posed in the case study, so from the file reading step to the final step of creating the results file. This code must be able to be rolled by your teacher without generating errors.

Evaluation

The evaluation will be based on the following criteria (20 points in total)

- Organization of your report (it should be clear, well structured and easy to read): /2
- Description of the methods used (a reader should be able to understand what you have done): /5
- Scientific accuracy of your analyses: /5
- Depth of your analyses: /5
- Results (realized profit) compared to a Benchmark model of professors and other teams in all sections of the course (English and French): /3