

Data Scientist

Practical Exercise

Introduction

These are the dimensions on which we will evaluate your work:

- Business understanding and storytelling
- Grasp of relevant methods and underlying concepts
- Critical thinking
- Technical acumen and tool proficiency

You are completely free to define your approach to the problem described below and add as much complexity to your solution as desired. Keep in mind however that we expect you to maximize the business value of your output vs the complexity of your approach at all times.

Practical Exercise

The data that will be provided to you comes (with modifications) from drivendata.org. They had a competition to predict from household surveys if a household is poor or not. You can find more context [here](https://drivendata.org).

Please use your machine learning skills and experience to develop a prediction model based on the data provided. We are not interested as much in the final result as in the process you follow, the insights you obtain, the machine learning workflow you design, and how you communicate your results and findings.

Data Scientist

Practical Exercise

Please keep in mind **all** the following milestones as part of your approach. They will all have the same weight on the final evaluation.

| | | |
|---|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Business Understanding | <ul style="list-style-type: none">• What is the problem?• What is available?• What does a solution look like? |
| 2 | Data Preparation | <ul style="list-style-type: none">• What tools will you use and why?• What steps are necessary?• Please include all developed code in your solution and document intermediate results |
| 3 | Model Building | <ul style="list-style-type: none">• What model(s) will you use and why?• How do you interpret and validate your model(s)? |
| 4 | Productization Plan | <ul style="list-style-type: none">• Once your model is ready how do you “deploy” it? |
| 5 | Storytelling | <ul style="list-style-type: none">• What are the main results/insights? What is the value of the model?• After receiving your solution we will invite you over to talk about it |

Instructions & next steps

Within a week of receiving the exercise, you'll send us back your report explaining the methods you used and your results including code, plots and anything else you find useful. If your exercise is successful, you will be asked to present your results together with your previous work to us at our offices in Madrid.

Thanks in advance!