Nuxtu SAS
Technical examination
Software Engineer - Machine Learning 13/Jun/2022
Deadline: 18/Jun/2022 - 23:59

For this technical examination you may use any books, notes, web pages, tutorials, documentation, or any other type of information source. You **may not** discuss this exam or questions related to the exam with your fellow contestants. It is also prohibited to ask for help from anyone you know that has experience on these topics, doing so will disqualify you from this and any future selection processes in *Nuxtu*.

Feel free to ask any question related to this examination to our *Chief Technology Officer* through his Nuxtu email oscar.diaz@nuxtu.co and CC ana.palomares@nuxtu.co
remember: **the only bad question is an unasked one**.

At the end of this examination, the following deliverables are expected:

- A forked GitHub repository from this original repository, with continuous commits on your progress. **Your last commit before the deadline will be evaluated**.
- A self-explanatory Jupyter Notebook with your solution for the problem, uploadedto the forked GitHub repository.
- Be prepared to answer any questions related to your solution. This will show us that you completely understand your implementation, which is the most important part of this examination.

Work hard. Have fun. Make history!

Jeff Bezos

The National Institute of Health (INS) has hired *Nuxtu* to develop a data science solution related to the COVID-19 global pandemic that has been afflicting our country for the last 2 years. The goal of the project is to create a predictive model that can determine the status of the patient (*Estado* column) by using as input any other columns or column combination that you see fit.

The INS has provided *Nuxtu* with the following public domain data, that can be used freely by our data scientist team:

- Data-set: https://tinyurl.com/nuxtu
- Docs: https://www.ins.gov.co/BibliotecaDigital/dataset-casos.pdf

Your assignment as the leader of the data science team is to complete the following tasks:

- Perform an Exploratory Data Analysis of the data-set, and deliver insights on your findings.
- Perform any data cleaning or pre-processing that you see fit.
- Deliver 2 different models that you consider suitable for the requirements of the prob- lem. **Remember: evaluate your models! Why would you prefer one model over another?**
- Get ready to tell a story to the *Nuxtu Board of Directors* that can help us decide if we should continue or not with the developing of this project.

The Jupyter Notebook that you deliver must be self-explanatory, so you must be sure that all required information is contained in the notebook. Do **not** upload the data-set to your GitHub repository.