

Clébio Marques de Oliveira Júnior

clebiomojunior@gmail.com | juniorcl.github.io | linkedin/in/clebiojunior | github.com/juniorcl

Introduction

I'm a physics teacher and a data scientist with a passion for technology. I currently develop projects that prioritize solving business problems, from understanding them, analyzing data, to extracting insights and implementing the solution. I also continue to develop myself with improvement and study activities such as a portfolio of data science projects and I also write about the same topic on a blog on [Medium](#).

For more details about my projects and each solution, they are described in my portfolio [juniorcl.github.io](#) or from my [GitHub](#) page.

Experiences

- **Data Scientist**, *Datarisk*, January 2022 - August 2024

My work is based on exploring, cleaning and analyzing data to bring results related to business problems with the help of machine learning. Below are some example projects:

- Python, Data Science, Machine Learning, Data Analysis, Git, SQL, Credit Score and Credit Risk.
- Development of a machine learning model with the purpose of determining a better credit score through the history of individuals.
- Development of a machine learning regression model for sales forecasting based on the history of previous years.
- Study and application of a clustering machine learning model for understanding different customer profiles.
- Implementation of a model that identifies certain customer behaviors, such as becoming premium, bad payer or having job instability.

- **Data Scientist**, *Be.XI*, March 2021 - January 2022

As a scientist I developed projects mainly focused on insights that can come from statistical analysis or machine learning models. Below there are some examples of these projects.

- Python, Data Science, Machine Learning, Data Analysis, Git and SQL.
- Treatment and preparation of texts for unstructured data acquisition through regular expressions (regex).
- Application of statistical methods to discover anomalous (outliers) points that can be harmful and fundamental for decision making.
- Implementation and use of machine learning models that can better describe the impacts of variables on the business.

- Machine learning model development to discover deliveries with the greatest possibility of delay through ranking metrics.

Education

- **Master's Degree in Natural Sciences**, *Universidade Estadual do Estadual do Norte Fluminense Darcy Ribeiro*, 2017 - 2019
- **Graduation, Natural Sciences - Degree in Physics**, *Instituto Federal de educação, ciência e tecnologia Fluminense*, 2010 - 2016

Complementary Courses

- **DS Comunity**, *Be A Data Scientist*, November 2022.
- **Data Science in Production**, *Be A Data Scientist*, November 2020.
- **Deep Learning with Python from Scratch**, *Udemy*, March 2020
- **Machine Learning and Data Science with Python from Scratch**, *Udemy*, October 2019

Skills and Knowledge

- SQL
- Python
- Data Analysis
- Machine Learning

Languages

- English, *Advanced*
- Esperanto, *Advanced*
- Portuguese, *Nativo*