

# Carlo Domenico Longo de Lemos

Sao Paulo, Brazil | [LinkedIn](#) | [GitHub](#) | [Machine Learning Blog](#)

---

## WORK EXPERIENCE

<b>Serasa Experian</b> (credit bureau)   Latin America DataLab	Sao Paulo, Brazil
<b>Staff Data Scientist</b> ( <i>Data Scientist Specialist II EB7</i> ) Full-time CLT	09/23 - Present
<ul style="list-style-type: none"><li>Since 04/23, I have led GenAI implementations across Serasa as the team's tech lead. Coordinating a team of four individual contributors in building NLP applications such as a legal copilot and a custom email campaign generator. Leveraged the use of retrieval augmented generation minimizing hallucination with inner source Python library for chatbots creation following guardrails.</li></ul>	
<b>Senior Data Scientist</b> ( <i>Data Scientist Specialist I EB8</i> ) Full-time CLT	05/22 - 08/23
<ul style="list-style-type: none"><li>Introduced a MLOps template using Docker and Airflow for automated client backtesting.</li><li>Developed an insurance claims forecasting framework that outperformed the previous model by over 10 ROCAUC points, surpassing industry benchmarks, generating 1.2M USD in new revenue in its first year.</li><li>Developed a causal inference model to estimate the optimal credit limit based on an individual's credit history.</li></ul>	
<b>Mid-level Data Scientist</b> ( <i>Data Scientist Analyst III EB9</i> ) Full-time CLT	11/21 - 04/22
<ul style="list-style-type: none"><li>Led a team of two junior data scientists in developing a credit score for the agricultural sector resulting in a product that raises 5M USD annually. Defined target population and designed business-centric metrics. Replaced external data source, improving model autonomy without sacrificing performance.</li></ul>	
<b>Itaú Unibanco</b> (bank)   Advanced Analytics Consulting Management	Sao Paulo, Brazil
<b>Junior Data Scientist</b> ( <i>Data Scientist I</i> ) Full-time CLT	11/20 - 10/21
<ul style="list-style-type: none"><li>Led a team of two data science interns in developing a model to set monthly sales targets to Itaú's bank officers using causal inference uplift/heterogeneous treatment effect models. A/B testing validated projected 3% annual revenue increase for tested products.</li></ul>	
<b>Data Science Intern</b> ( <i>Intern</i> ) Paid internship	10/19 - 10/20
<ul style="list-style-type: none"><li>Designed causal inference model for pricing, assessing Itaú portfolio elasticity. Introduced a new pricing dimension, impacting 30M customers.</li><li>Enhanced business processes using pandas, SQL and Hive, reducing operation time.</li></ul>	

## EDUCATION

<b>University of São Paulo (USP)</b>	08/21 - 07/24
Master's degree in Probability and Statistics - Threshold tuning of the probability simplex in multiclass classification <a href="#">GPA 4/4</a>	
<b>Federal University of ABC (UFABC)</b>	05/15 - 04/21
Bachelor's degree in Science and Technology <a href="#">GPA 3.88/4</a>	

## OPEN KNOWLEDGE AND OPEN SOURCE SOFTWARE

<b>Machine Learning blog</b>	<a href="#">vitaliset.github.io</a>
Vitali Set is my personal site where I share insights and knowledge about Machine Learning.	
<b>scikit-learn</b>	<a href="#">All my pull requests to scikit-learn</a>
<b>datalib</b>	<a href="#">All my pull requests to datalib</a>

## TOOLS I USE DAILY

functional and oop with python, numpy, matplotlib, scikit-learn, pytorch, tensorflow, keras, xgboost, langchain, pandas, sql, pyspark, git, pytest, ci/cd, airflow, docker, azure, linux.

---

## RESEARCH FELLOWSHIPS

<b>Wada Property in Doubly Transient Chaos</b>	11/18 - 09/19
FAPESP (São Paulo Research Foundation)	
<b>The Euler-Maclaurin formula and a few applications in Pure and Applied Mathematics</b>	09/17 - 08/18
FAPESP (São Paulo Research Foundation)	
<b>Numerical study of orbits in the 3-body problem</b>	08/16 - 07/17
CNPQ (Brazilian National Council for Research and Technological Development)	
<b>Introduction to stellar structure and evolution</b>	10/15 - 07/16
UFABC, Researching Since the First Day Program	