# MARUAN BAKRI OTTONI

📱 +55 41 99209-2810    ✉ [maruanbakriottoni@gmail.com](mailto:maruanbakriottoni@gmail.com)    ⚡ [Github](#)    🖥 [mbottoni.github.io](#)

## EDUCATION

**MSc in Computer Science**, University of São Paulo (USP)                                    *Ongoing*
Research focus on computational neuroscience and random graph modeling of neural connections and social networks
*Coursework*: Graph Theory, Statistical Methods, Deep Learning, NLP, Parallel Computing

**BSc in Electrical Engineering**, University of Campinas (Unicamp), Brazil            *2018 – 2022*
GPA: 8.5/10
*MSc/PhD level coursework*: Information Theory, Stochastic Processes, Bayesian Inference, Machine Learning
*Research Projects*: Introduction to Lattice-based Cryptography(CNPQ Scholarship 2021-2022), Information Geometry for Machine Learning (CNPQ Scholarship 2022-2023)

## PROFESSIONAL EXPERIENCE

**Data Scientist**                                                            September 2023 - Present
CloudWalk Inc.                                                                        *São Paulo, Brazil*
As part of an exploration team initiative my job was to propose new state of the art ideas to improve the AI usage across the company. These include:

- Designed and deployed scalable monitoring tools to track the health of over 3000 ML features in our feature store, improving model reliability and observability.
- Built generative diffusion models for synthetic data generation on the context of tabular features for downstream tasks such as credit scoring and fraud detection, helping the improvement of classification metrics for the models in about 1%.
- Pretrained and finetuned foundational model for credit assignment using LLM inspired architectures, enabling the deployment of credit scoring solutions that led to over R\$ 100 million in increase of approved credit for clients.
- Led end-to-end development of experimental AI projects involving mostly self-supervised learning, graph-based learning (GNN's) for fraud detection, and mechanistic interpretability for large language models (LLMs).

**Quant Analyst Summer Internship**                                      January 2023 - March 2023
Goldman Sachs                                                                        *São Paulo, Brazil*
As part of a quant team I helped the development of a linear optimization software for asset inventory management across different entities of the bank

- Engineered algorithmic solutions that optimized asset allocation and significantly reduced redundant asset movements, improving operational efficiency.

**Quant Analyst Internship**                                          October 2022 - November 2022
UHedge Trading Solutions                                                              *Campinas, Brazil*

- Conducted mathematical analysis of derivatives strategies to evaluate trading opportunities and risk-return profiles

## RELEVANT SIDE PROJECTS

**Logit Graph** Developed a library for probabilistic graph generation via a logit-based edge process. Did also iterative benchmarks on current statistical graph models, achieving state of art performance for the method developed. [Lib](#)

**Horror Movie Recommendation** Built a FastAPI app that recommends horror movies from OMDb based on user prompts using a unified retrieval pipeline: Sentence-Transformers embeddings with blended priors (IMDb rating/votes, facet overlap, recency), optional cross-encoder reranking. Added UI, PostgreSQL, CI/CD, Payments integrations. [Site](#), [Code](#)

**RAG Solution for Analyzing Auction Documents** Developed a Retrieval-Augmented Generation (RAG) chatbot using Langchain and MongoDB, with a chat interface built in Flask. The bot enables interaction with potentially hundreds or thousands of documents across multiple auctions, allowing users to efficiently navigate and retrieve relevant information. [Code](#)

**Graph Similarity via Representation Learning** Contributed with ML methods for graph distance, implementing SOTA architectures and matching strong baselines. [Code](#)

## SKILLS & LANGUAGES

| | |
|---|---|
| **Programming** | Python, Matlab, C, Java |
| **Tools** | Google Cloud Platform, Docker, Git, Linux, Kubernetes, Spark |
| **ML/AI** | PyTorch, TensorFlow, Scikit-learn, FSDP Training, Hugging Face |
| **Data/Viz** | SQL, Matplotlib, Seaborn, Plotly |
| **Languages** | Portuguese (Native), English (Fluent C2), French (Basic A2) |