

# MARCO H A INÁCIO

Doctor in Statistics & Machine Learning Researcher

@ marcoinacio.com   marcoinacio.com   randommm   Born in 1991



## EDUCATION

- **PhD in Statistics** (University of São Paulo and UFSCar) concluded in August 2020. Was a visiting researcher at BME (Budapest University of Technology and Economics). Research emphasis in machine learning
- **Master in Statistics** (University of São Paulo and UFSCar) concluded in May 2017. Research emphasis in Bayesian statistics.
- **Bachelor in Economics** (University of São Paulo). Research emphasis in car insurance modeling.

## JOURNAL ARTICLES

- Distance assessment and analysis of high-dimensional samples using variational autoencoders, Information Sciences (Marco Inácio, Rafael Izbicki, Bálint Gyires-Tóth). DOI 10.1016/j.ins.2020.06.065.
- The NN-Stacking: Feature weighted linear stacking through neural networks, Neurocomputing (Victor Coscrato, Marco Henrique de Almeida Inácio and Rafael Izbicki). DOI 10.1016/j.neucom.2020.02.073.
- What if the forecaster knew: assessing forecast reliability via simulation, Trends in Computational and Applied Mathematics.
- Bayesian superposition of pure-birth destructive cure processes for tumor latency, Communications in Statistics - Simulation and Computation (Josemar Rodrigues, Marco Henrique de Almeida Inácio, Adriano K. Suzuki, Fernando Raimundo da Silva and Narayanaswamy Balakrishnan). DOI 10.1080/03610918.2018.1538455.
- Comparing two populations using Bayesian Fourier series density estimation, Communications in Statistics - Simulation and Computation (Marco Henrique de Almeida Inácio, Rafael Izbicki and Luis Ernesto Salazar). DOI 10.1080/03610918.2018.1484480.

## INT'L EXPERIENCE

- Hungary: from March 2019 to June 2019; from August 2019 to February 2020; from August 2020 to November 2020. Purpose: research at BME, with emphasis on machine learning, deep learning, sequential data modeling and variational autoencoders.
- New York, United States: February 2016. Purpose: visit Stan research group at Columbia University.
- Montreal, Canada: from December 2012 to January 2013. Purpose: study French.

## EXPERTISE

- Fields of expertise: data science, machine learning, analysis, Statistics, neural networks/deep learning, Bayesian inference.
- Languages: Portuguese (native), **English (C1)**, Hungarian (B1/B2), French (A1).
- Programming languages experience (in decreasing order): **Python, Stan, R, C++, HTML, CSS, JavaScript, Bash, SQL, Arduino, PHP, Ruby.**
- Other tools: **Stan, Pytorch, Git/Github/Gitlab, Conda, Docker, SSH, Django, Google Cloud, Pyro, PySpark, Google Cloud, network configuration, Linux server administration.**
- Member of Stan development team: <https://mc-stan.org/about/team/>, a **Bayesian Statistical** framework started by Columbia University written in C++.
- Introductory Python and machine learning tutorial: <https://pytutorial.marcoinacio.com>.

## PRACTICAL EXPERIENCE

- Since Feb 2021, working as a data scientist at Boa Vista SCPC.
- Work with **Pytorch**, sequential data, **LSTMs** and other deep learning models at BME (Budapest University of Technology and Economics), in particular, with text data.
- Contributor of **Stan**, using **C++**.
- Applied work with **car insurance**, Bayesian Statistics, Stan and zero-inflated models during undergraduation.
- Worked with **variational autoencoders**, CIFAR10 images and two-sample comparison (published on high impact journal).
- Worked with ensembling of machine learning regression models using neural networks (published on high impact journal).
- Work with **interpretable machine learning neural networks** (in progress).
- Development and maintenance of website <https://fifaexperts.com> using **Python, Django, JQuery/javascript, CSS, SQLite and HTML.**
- Interactive applications with Dash, e.g.: <https://covid.marcoinacio.com/>.
- Delivered web and GUI applications for a company in the past.

## PREPRINTS

---

- arXiv:2004.14479. Monte Carlo simulation studies on Python using the *sstudy* package with SQL databases as storage. Marco H A Inácio.
- arXiv:1910.05206. NLS: an accurate and yet easy-to-interpret regression method. Víctor Coscrato, Marco Henrique de Almeida Inácio, Tiago Botari, Rafael Izbicki.
- arXiv:1908.00105. Conditional independence testing: a predictive perspective. Marco Henrique de Almeida Inácio, Rafael Izbicki, Rafael Bassi Stern.

## ... PAST GRANTS

---

- Brazilian undergraduate research grant from CNPq. 2012-2013.
- Brazilian Master grant from CAPES. 2015-2017.
- Brazilian PhD grant from CAPES. 2017-2020.
- Erasmus PhD research exchange (to Hungary) grant. 2019.
- Brazilian PhD research exchange (to Hungary) grant from CAPES. 2019-2020.
- Hungarian BME-Artificial Intelligence FIKP grant of Ministry of Human Resources (BME FIKP-MI/SC). 2019-2020.

## ... TEACHING

---

- 20 hours class of cloud computing at Metropolitan University of Campinas. 2021.
- 60 hours trainee teacher for introduction to probability and data mining at Federal University of São Carlos. 2017-2018.
- Short 2 or 4 hours courses on Stan. 2016-2017.
- Weekly 1 hour lectures on Machine Learning/Statistics at Boa Vista SCPC. 2021.

## ... CONFERENCES

---

- Conditional density estimation using Fourier series and neural networks. Marco Inácio, Rafael Izbicki. Symposium on Knowledge Discovery, Mining and Learning (KDMiLe), 2018.
- A Bayesian approach to density estimation via Fourier orthogonal series. Marco Henrique de Almeida Inacio, Rafael Izbicki, Luis Ernesto Salasar. Simpósio Nacional de Probabilidade e Estatística (SINAPE), 2016.
- Aplicação de modelos de regressão inflacionados de zero para o mercado brasileiro de seguros de automóveis. Marco Henrique de Almeida Inácio and Bruno Cesar Aurichio Ledo. Simpósio Internacional de Iniciação Científica e Tecnológica da Universidade de São Paulo (SI-ICUSP), 2013.

## ... OTHER WORKS

---

- Introductory Python tutorial, focused specially on machine learning for statisticians: <https://pytutorial.marcoinacio.com>.
- *sstudy*: a Python package designed to simplify the preparation of simulation studies using SQL database engines as the storage system: <https://sstudy.marcoinacio.com>.
- Introduction to Stan with Python and R source codes (in Portuguese): <https://marcoinacio.com/stan>.
- Designed Worldcup probabilities competition website: <https://fifaexperts.com/>. A paper on our research has been accepted at Trends in Computational and Applied Mathematics.