Bruno Gomes Coelho

Last updated July 2021

bruno.coelho@nyu.edu
brunogomescoelho.github.io
in/bruno-coelho-277519129/

EDUCATION

New York University (NYU) — Brooklyn, NY, USA
PhD in Computer Science focused on Machine Learning and Visualisation — *Set. 2020 – Jun. 2025 (predicted)*

- **Relevant classes**: Probability for Data Science, Machine Learning, Deep Learning, Information Visualization

University of São Paulo (USP), ICMC — Campus of São Carlos, Brasil
Bachelor's degree in Computer Science; Emphasis in Data Science; GPA: 8.8/10.0 — *Fev. 2016 – Jan. 2020*

WORK EXPERIENCE

University of São Paulo, ICMC — São Carlos, São Paulo
Data Science MBA T.A. — *Feb 2020 -*

- **Processing**: Creating classes, practical exercises/evaluations and tutoring a class of 150 students.
- **Classes mentored**: Python programming; Intro to ML; ML; Advanced data acquisition and processing; Neural networks and applications

Consulting with The AI Academy — Remote
Data Scientist — *May 2020 - Oct 2020*

- **Continuous Development for DS**: Prototyping and evaluating new DS processes and tools for use in the finance sector with client being one of Brazil's biggest banks.
- **DevOps**: Integrating H2O Driverless AI into automated CD pipelines.

Imaflora Origens — São Carlos, São Paulo
Data Scientist — *Feb 2020 - July 2020*

- **ETL**: Data wrangling business and Geo-spatial data from different sources for a interactive visualisation.
- **Visualisation**: Building a interactive platform for exploring new data-driven insights. Working alongside business owners to visualise and understand the impact of social programmes.

University of São Paulo, ICMC — São Carlos, São Paulo
Research Assistant - international scholarship by USAID — *Mar 2017 - Dez 2019*

- **Processing**: Transformed raw audio data (mosquito's wingbeats) into structured data using time series analysis.
- **Classifying**: Evaluated different models (tree classifiers, deep neural networks, SVMs, etc.) to classify different insect species; achieved highest classification accuracy (99%) in our yearly milestone report.
- **Undergratuate thesis**: Analysing CNNs applied directly to raw audio data for insect classification; Demonstrating the problems with classical classification approaches in the context of sensor generalization.

Getúlio Vargas Fundation (FGV) and Imaflora — São Carlos, São Paulo
Data Scientist — *Aug 2019 - Dez 2019*

- **Data Processing**: Creating an open source library for processing Brazil's timber exploration
- **Illegal Exploration**: Developing a django/D3 visualisation tool that integrates different systems and facilitates the discovery of anomalous data. Integration of 5 years of data across 3 government maintained systems.

Birdie — São Carlos, São Paulo
Data Science Intern — *Aug 2018 - Fev 2019*

- **Data Acquisition and Mining**: Acquired more than 1 million e-commerce products through crawlers, building a reliable back-end infrastructure through Scrapy and Airflow; Extraction of useful features from the unstructured textual data of the products.
- **Machine Learning**: Created hierarchical models for product category classification using transductive ML models; Deployed the models as micro-services for processing 10 million+ daily products.

Major Projects

- **Winner of the Dotz NLP Competition**   May 2020 - July 2020

  - ○ **Competition**: Hierarchical classification of e-commerce products from their description.
  - ○ **Solution**: Building a integrated pipeline using a mixture of portuguese Bert, Character CNNs, traditional sparse BOW model and hierarchical specific models.

- **Finalist of the Data Science Games 2018**   May 2018 - Set 2018

  - ○ **Qualification Phase - Kaggle competition**: Analysed the bond market with data provided by BNP Paribas; Predicted customer interest in a bond using diverse data, from customer data to international market information; Final result in the top 30 worldwide from 220 participating universities, guarantying a place in the final.
  - ○ **Final - 2 day hackaton in France**: Used e-commerce data to help predict customer interest in online shopping; Achieved original company's baseline in less than 30 hours. Team sponsored by Itaú Unibanco bank.

- **Co-founder of DATA - Extracurricular Data Science Group**   September 2018 - Jan 2020

  - ○ **Co-Founder**: Creating and structuring a extracurricular data science group recognised by the university
  - ○ **Teaching**: Weekly classes teaching basic Machine Learning concepts (statistical learning theory, K-NN, Linear Regression, SVMs, etc) and frameworks (numpy, pandas, seaborn, scikit-learn) to a class of 30 students.
  - ○ **Deep Learning**: Head of weekly study group focused on state of the art image recognition papers. Development of practical projects with financing from Santander bank.

Papers

- Optimal Team Economic Decisions in Counter-Strike - Peter Xenopoulos, **Bruno Coelho** and Claudio Silva, AISA-2021
- NE-Motion: Visual Analysis of Stroke Patients Using Motion Sensor Networks - RC Contreras, A Parnandi, **BG Coelho**, C Silva, H Schambra, LG Nonato - MDPI Sensors 2021

Technical Skills

Languages: C, Java, Python, JavaScript, SQL
DS/ML/Viz: SciPy, Scikit-Learn, Pandas, Jupyter, PyTorch, TensorFlow, DVC, Seaborn, D3, Plotly Dash
Others: BeautifulSoup, Scrapy, Django, Git, Docker, Airflow, Driverless AI

Awards

- Winner of the Dotz NLP Competition - Alongside a team member, achieved first place building hierarchical Portuguese NLP models from 80 participating teams.
- Finalist of the Data Science Games 2018 - one of the 20 best teams, from 400 participating teams, representing 220 universities across 46 countries.
- Silver medal on Elo's Merchant Category Recommendation on Kaggle. Over 4000 teams of 1-4 people took part.
- 19th place at Data Mining Cup, from 149 Teams from 114 universities in 28 countries - Best team from Brazil.
- Honorable mention on best article at USP's International Symposium of Undergraduate Research - International Stage. Less than 600 mentions given from over 3200 applicants.
- Recipient of Academic Featured Student (2016), for students with a GPA above 9.0/10.0 at USP. Around 7 in each 120 students are awarded each year.

Teaching

Python for Data Science course - Invited teacher   Setember 2019 - December 2019

- **Teaching**: Teaching fundamentals of Python to industry members. Course given to 80 students alongside 3 other teachers. Official course at the Universidade de São Paulo.

Invited Data Science Course Lead   Setember 2019 - December 2019

- **Teaching**: Heading the Data Science part of the Data Science and Engineering course at a startup incubator. Teaching 40 students about topics such as Pandas, Matplotlib, Scikit-Learn, Introductory to Machine Learning, Boosting and Flask.

Invited Events

- **Khipu**: Latin American Meeting In Artificial Intelligence, Uruguai, invited participant through full merit-based scholarship
- **Data Day**: Head of the discussion panel "Present and Future of AI"with invited professors and industry members.
- **5th Mathematics Solutions Workshop for Industrial Problems**: One of the 3 invited undergraduate members. Analysing timber transport data for an ONG.
- **HTN 2018**: Full sponsorship to Hack The North 2018, Canada's biggest hackaton. Developed InstaFace, a project to make use of image recognition to find friends on the street.

Voluntering

- **Programming Tutor (12 months)**: Taught fundamentals of logic and computer programming to underprivileged communities, through the basics of web development, integrating JavaScript code with HTML and CSS pages.
- **Tutor (8 months)**: Personal mentor to high school student, helped with doubts related to university life, building a study plan for the SATs and answering carreer questions.