Bruno Gomes Coelho

Last updated January 2023

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 — *September 2020 – Expected June 2025*
  - **Relevant classes**: Probability for DS, ML, DL, Algorithmic ML and DS

- **New York University (NYU)** — Brooklyn, NY, USA
  Master's in Computer Science — *September. 2020 – May 2022*

- **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 — *February 2016 – January 2020*

INDUSTRY EXPERIENCE

- **The AI Academy** — Remote
  Data Scientist Consultant — *May 2020 - October 2020*
  - **Continuous Development for DS**: Prototed and evaluated new DS processes and tools for use in the finance sector with client being one of Brazil's biggest banks.
  - **DevOps**: Integrated H2O Driverless AI into automated CD pipelines.

- **Imaflora Origens** — Brazil
  Data Scientist — *February 2020 - July 2020*
  - **ETL**: Data wrangling business and Geo-spatial data from different sources for a interactive visualisation.
  - **Visualisation**: Built a interactive platform for exploring new data-driven insights. Worked alongside business owners to visualise and understand the impact of social programmes.

- **University of São Paulo, ICMC** — Brazil
  Research Assistant - International scholarship by USAID — *March 2017 - December 2019*
  - **Processing**: Transformed raw audio data (mosquito's wingbeats) into structured data using time series analysis.
  - **Classifying**: Building pipelines and evaluating models to classify different insect species.
  - **Undergratuate thesis**: Applying CNNs to raw audio data for insect classification; Analysing covariate shift in the context of sensor generalization.

- **Getúlio Vargas Fundation (FGV) and Imaflora** — Brazil
  Data Scientist — *August 2019 - December 2019*
  - **Data Processing**: Integration of multiple government systems related to timber exploration data.
  - **Illegal Exploration**: Developing a django/D3 visualisation tool to facilitate the discovery of anomalous data.

- **Birdie** — Brazil
  Data Science Intern — *August 2018 - February 2019*
  - **Data Acquisition and Mining**: Acquired more than 1 million e-commerce products through crawlers, building a reliable back-end infrastructure; Extraction of useful features from the unstructured textual data of the products.
  - **Machine Learning**: Building hierarchical models for product category classification
  - **Production**: Deployed the models as micro-services 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 - September 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**: Used e-commerce data to help predict customer interest in online shopping.

- **Co-founder of DATA - Extracurricular Data Science Group**   September 2018 - January 2020
  - **Co-Founder**: Creating and running a extracurricular data science group recognised by the university.
  - **Teaching**: Weekly classes teaching basic Machine Learning concepts and frameworks to a class of 30 students.

Technical Skills

Languages: Python, SQL, C, JavaScript, Java
Topics: NLP, Computer Vision, Deep Learning, MLOps, Big data
DS/ML: SciPy, Scikit-Learn, Pandas, Jupyter, PyTorch, TensorFlow, DVC, Seaborn, D3, Plotly Dash
Others: BeautifulSoup, Scrapy, Django, Git, Docker, Airflow, Driverless AI, GCP

Awards

- Winner of the NDD Tech Challenge - First place building ML models to classify incomplete fiscal documents
- 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 / 10. Awarded to 7 / 120 students.

Teaching

- **Data Science MBA T.A.**   February 2020 - current
  - **Classes (150 students):**: Intro to Python; Intro to ML; Advanced data acquisition and processing; ML; DL

- **General Computer Science and Data Science**
  - **NYU's Introduction to CS**: Python programming general grading, Fall 2021.
  - **Invited DS Course Lead**: Heading the DS part of a 3 month course at a startup incubator, Summer 2019.
  - **Python for DS course**: Teaching industry members at the Universidade de São Paulo, Fall 2019.

Invited Events

- **ACM's Internet Measurement Conference 2022**: Invited participant and travel grant winner
- **Khipu 2019**: Latin American Meeting In Artificial Intelligence, invited participant with 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.

Voluntering

- **Programming Tutor, 2017, 1 year:** Taught fundamentals of computer programming to underprivileged communities.
- **Tutor, 2018, 8 months:** Personal mentor to high school student, helping with SATs and career questions.