

# LUIGI NOTO

21 Stuyvesant Oval Apt. 6G, New York, NY, 10009

+1 (917) 582-2478 · ln2205@nyu.edu · linkedin.com/in/luiginoto · luiginoto.github.io

## EDUCATION

---

**New York University**, New York, NY

Expected 05/2023

*Master of Science in Data Science*, GPA: 4.0/4.0

- Relevant coursework: Probability and Statistics, Optimization and Linear Algebra, Machine Learning, Big Data, Information Theory

**Bocconi University**, Milan, Italy

07/2021

*Bachelor of Science in Economics, Management and Computer Science*, Grade: 110/110 cum laude

- Relevant coursework: Advanced Mathematics and Statistics, Computer Programming, Machine Learning, Decision and Game Theory, Big Data and Databases
- Bachelor Thesis: *Bayesian Learning: Clustering with the Dirichlet Process Mixture Model*
- Supervisor: Prof. A. Lijoi

## EXPERIENCE

---

**Bocconi University**, Milan, Italy

02/2021 - 08/2021

Undergraduate Teaching Assistant

- Contributed to the improvement and typesetting in LaTeX of about 50% of teaching material for the bachelor class in Applied Stochastic Processes.
- Proofreading of faculty lecture notes to ensure consistency in notation for later publication.

**Barclays**, Knutsford, UK

07/2020 - 08/2020

Technology Intern Data Analyst - Barclays UK Infrastructure Estate Management

- Gathered, analyzed and summarized data about Total Cost of Ownership and Technical Debt of Barclays Technology Estate, through the consolidation of multiple monthly feeds to highlight trends and direct cost optimization initiatives, using SQL and data analysis software.
- Implemented data validation checks and automated data collection and concatenation tasks with Excel VBA macros, saving 3-5 hours per week of manual work to consolidate data and resolve data integrity issues.

## SELECTED PROJECTS

---

**MovieLens Recommender System and Extensions.** Built and evaluated a collaborative-filtering recommender system with Spark alternating least squares (ALS) implementation using the MovieLens dataset. Compared Spark's parallel ALS model to lenskit single-machine implementation in terms of efficiency and model performance. Implemented accelerated search at query time using Spotify's annoy spatial data structure.

**Analysis of Weather Data in Australia.** Performed hypothesis testing (resampling methods) on temperature data to investigate climate change. Trained and validated ML classification models to predict next day rain occurrence (XGBoost with 0.90 test AUC, 0.75 test average precision score).

**Loan Default Prediction.** Trained and validated ML classification models to predict the probability of default on a loan with a dataset about US small businesses provided by the US Small Business Administration (SBA), available on Kaggle (XGBoost with 0.93 test accuracy, 0.93 test sensitivity).

## SKILLS

---

**Computer** Python (Numpy, Pandas, SciPy, Matplotlib, Scikit-learn, xgboost, PyTorch), Latex, R, SQL, Spark (PySpark), Git

**Language** Italian (Native), English (Fluent), French (Basic)

## CERTIFICATIONS

---

- Neural Networks and Deep Learning – Coursera

## HONORS AND ACTIVITIES

---

- 3rd place at Vodafone Data Science Hackathon out of 40 participating teams (2021)
- Bocconi Merit Award, awarded to outstanding students based on academic standing (2020)
- Member of *Bocconi Students for Data Science* student association (09/2019-07/2021)

I hereby authorize the treatment of my personal data according to the local law.