

# Davide Vaccari

[dv2438@columbia.edu](mailto:dv2438@columbia.edu) | <https://www.linkedin.com/in/davidev7/> | <https://github.com/davidev7>

## EDUCATION

---

### Columbia University – New York City, U.S.A.

M.A. Quantitative Methods in the Social Sciences - Major: Data Science, Minor: Finance

Sep 2020 - Oct 2021

Current GPA: 4.1/4.0

#### Thesis:

Analyzed retail investors' roles in stocks' price swings during fads.

- Adapted traditional forecasting models with proxies of retail investors' attention such as Google searches for tickers and discussion in forums (Reddit);
- Built a "wide & deep" model (ensemble ARIMA and LSTM for time series) to assess how much these proxies add to prices forecasts.

#### Projects:

Won "Predict World Happiness Rankings" hackathon organized by Columbia's statistics department.

[Link to my notebook.](#)

Won "Covid RX" hackathon organized by Columbia's statistics department.

- Developed a ML algorithm to predict whether an individual's RX was healthy or had signs of either Pneumonia or Covid-19, achieving an accuracy of 99.2%. Project for Columbia's health department.

[Link to my notebook.](#)

#### Exams:

Data Analysis and Econometrics	A	
Theory and Methodology for Quantitative Research	A+	
Machine Learning	A	
Projects in Advanced Machine Learning	A+	
Modern Data Structures	A+	
Data Visualization	A	<a href="#">Portfolio of my visualizations.</a>
Corporate Finance	B+	
Behavioral Finance	Pass	
Natural Language Processing	NRA student	

### Università degli Studi di Ferrara – Ferrara, Italy

B.A. Economics

Sep 2016 - Oct 2019

Grade: 100/110

#### Thesis:

"China reshaping Mediterranean trade."

Developed a gravity model to forecast the economic impacts of the "Budapest - Piraeus" railway, once finished, on European and Mediterranean Trade. The railway is a core part of the Belt & Road Initiative by the Chinese government.

- Collected estimated costs and budgets of the railway sections (crosses 4 countries);
- Calculated the costs until completion considering political risks;
- Forecasted the expected time reductions for trade given capacity and policies;
- Estimated the increase in trade in the countries involved given the construction of the railway.

#### Exams:

Computer Science	30 cum laude
Political Economy	30/30
Public Economics	30/30
Financial Mathematics	30/30
Development Economics	30/30
Economic History	30/30
Microeconomics	27/30

## PROFESSIONAL EXPERIENCE

---

### Fondazione Eni Enrico Mattei – Milan, Italy

Sep 2021 - Present

Research Associate

Managing CITY 4 GREEN European Union's project:

- Applying a quantitative and qualitative model to evaluate environmental impacts of public investments in Central Italy;
- Outline insightful findings with data visualizations.

### Confindustria Emilia – Modena, Italy

Apr 2020 - Jul 2020

External Consultant - "Eyewitness Project"

- Designed a ML algorithm for the recognition of retinal diseases. Achieved 70% accuracy with a dataset of just 7,000 images on whether an eye was healthy or not.
- Led the market research and the search for partners.
  - Defined the contract with the head of the Italian ophthalmologist association to include their database in the study;
  - Introduced Confindustria to the final partner for hardware implementation.

### Fondazione Eni Enrico Mattei – Milan, Italy

May 2019 - Jul 2019

Research Assistant

- Mined data from the National Institute of Statistics for the "SDSN Italia SDGs City Index 2020";
- Co-authored a white paper on Green Finance.

### Smart Farming, Startup – Milan, Italy

Apr 2018 - Nov 2018

Founder

Founder of a startup that didn't take off. The hardware wasn't ready for market deployment. So over time, everyone preferred to pursue their own career. However, it was an important and valuable experience. We aimed to optimize resources in the agricultural field thanks to the use of sensors and ML algorithms.

- Presented the project at the Sixth International Conference on Sustainable Development at Columbia University and at the SDSN world meeting in New York City.

## COMPUTER SCIENCE SKILLS

---

**Machine Learning:** Supervised learning methods, including OLS models, linear models, support vector machines, decision trees and random forests, and gradient boosting, both for regression and classification [**Scikit-learn**]. Calibration, model evaluation and strategies for dealing with imbalanced datasets [**Pandas, Numpy**]. Unsupervised techniques: PCA, clustering and cluster evaluation, manifold learning. Advanced machine learning models, including Deep Learning, Convolutional Neural Networks for image and text data, Object detection models, Recurrent Neural Networks (Time-series data), and Adversarial Neural Networks [**Tensorflow, Keras**]. Natural Language Processing [**NLTK**].

**Data Visualizations:** Plots, Graphs, Maps, and word clouds [**ggplot2, ggraph, ggmap**]. Both interactive and static. An [example](#).

**Software:** Python, R, MySQL, Office tools (Word, Excel, PowerPoint) and Mac/Google equivalents.

## CERTIFICATIONS

---

Artificial Intelligence & Machine Learning

Bloomberg Terminal [2021]

- LinkedIn tested Badge (top 30% out of 241k test takers) [2021]

IELTS: 7.5/9 [May 2018]

GRE: 321/340 [Dec 2019]

- Confindustria - Corporate Training [Jan - Jul 2020]

- Datacamp [2020/2021]

## LANGUAGES

---

English - Full professional proficiency

Italian - Native proficiency