

# Mark Ibrahim

mark.ibhm@gmail.com • (973) 459-8429 • [markibrahim.me](mailto:markibrahim.me)  
in [marksibrahim](#) • [github.com/marksibrahim](https://github.com/marksibrahim) • [scholar](#)

## EXPERIENCE

New York, NY

### Staff Research Engineer, Fundamental Artificial Intelligence Research (FAIR), Meta AI

July 2019 – Present

- Researching methods to measure and improve the common sense of deep learning vision/language models
  - Designing generalization mechanisms in self-supervised learning, led to a +10% in classification robustness
  - Authoring research papers (e.g., NeurIPS, ICLR, CVPR) to [explain and structure](#) representations
- Architected tooling to track and launch experiments with distributed multi-GPU training in PyTorch
- Manager for interns/AI Resident; instructor for Georgia Tech's Deep Learning course (+10k students)
- Led creation of [Ridge Sketch](#) library for large-scale ridge regression and contributor to [CrypTen](#)

### Senior Machine Learning Engineer, Center for Machine Learning at Capital One

Sep 2016 – June 2019

- Led Explainable AI team to build tools and research for explaining black-box deep learning models
  - Built open-source Python [library](#) to generate global explanations for neural network predictions
  - Published 2 interpretability research papers (NeurIPS workshop 2018 and ACM AAAI 2019)
- Engineered a real-time [notification](#) system for predicting mistaken charges on 10 million transactions per day
- Developed deep learning (RNN + LDA) customer archetype model in collaboration with Columbia Prof. John Paisley

### Data Engineering Fellow (2016) & Technical Advisor, Insight Data Science

May 2016 – June 2018

- Developed a graph-based knowledge search engine ([knowledgegraph.us](#)) powered by Wikipedia
  - Distributed parsing of all 5 million articles using Spark on Amazon Web Services (AWS)
- Designed a D3.js user interface powered by a graph database (Neo4j), Elasticsearch, and Python (Flask)

### Quantitative Portfolio Risk Analyst, UBS

Jun 2012 – Aug 2014

- Applied unsupervised machine learning (PCA) to identify \$570k in uncaptured sensitivity to 0.01% move in rates
- Automated daily 2½ hour manual risk calculation for \$658 million trading portfolio in Python

## SELECT RESEARCH

“ImageNet-X: understanding model mistakes with factor annotations”—*BY Idrissi, M Ibrahim. ICLR Spotlight 2023.*

“Grounding inductive biases: invariance stems from data”—*D Bouchacourt, M Ibrahim, A Morcos. NeurIPS 2021.*

“Global Explanations of Neural Networks: Mapping the Landscape of Predictions”—*M Ibrahim et al. AAAI 2019.*

“Connecting Every Bit of Knowledge: [Wikipedia's First Link Network](#)”—*M Ibrahim et al. J. Computational Science 2017*

**Patent:** “Techniques to perform global attribution insights in neural networks”—*US Patent 16/855,685*

**Select Talks:** ICLR Spotlight (2023), NeurIPS Spotlight (2021), PyCon (2020), AAAI Spotlight Talk (2019), NYC Python Meetup (2018), Tom Tom Machine Learning Conf (2018), [Data Driven](#) at George Washington U. (2017).

## COMMUNITY

**Researcher** for the [AI-Powered](#) COVID-19 Forecasting [Data for Good](#) Program with Direct Relief non-profit org  
**Co-A** for Columbia U. Data Science Masters Capstone (2019-2018). **Co-organizer** Vermont Python User Group (2016)

**Mentor** for NYU PhD Student (FAIR/NYU; 2022-present) and Columbia U. Data Science Masters Capstone (2019-2018)

## EDUCATION

**Statistics MicroMasters**, Massachusetts Institute of Technology (MIT)

**M.S. Mathematics**, University of Vermont

Course Instructor: [Calculus I](#) (72 students) and [Calculus II](#) (38 students)

**Honors B.A. Mathematics**, *Magna Cum Laude*, Hamilton College

[19th Gold Scholar](#) for student of “highest standards.” *Phi Sigma Iota*: highest honor for foreign languages