Miguel Rivera-Lanas

rivlanm@protonmail.com | 703-981-0554 | https://www.linkedin.com/in/mriveralanas/

EDUCATION

Carnegie Mellon University - Pittsburgh, PA **GPA: 3.8**
Master of Science: Data Science and Public Policy **May 2025 (Expected)**

- **Relevant Coursework**: Machine Learning: Intro to Machine Learning (ML Dept.), Intro to Deep Learning (ML Dept.) Public Policy: Policies of Wireless Systems, Federal Budget Policy, Management: Optimization, Decision & Risk Modeling

University of Pennsylvania - Philadelphia, PA **GPA: 3.5**
Bachelor of Arts: Economics, Minor in Statistics, International Relations **May 2017**

PROFESSIONAL EXPERIENCE

Allegheny County Dept. Health & Human Services (DHS) **Pittsburgh, PA**
Data Science Intern - Office of Technology & Analytics **November 2024 - Present**

- Developed ML models to identify individuals most at risk of facing homelessness in Allegheny County.
- Improved Random Forest precision at K=100 via feature and hyperparameter engineering.
- Developed fairness assessments (FPR and predictive parity) across demographic categories.

Carnegie Mellon University **Pittsburgh, PA**
Research Assistant – Privacy Engineering **Jan. 2024 – Sept 2024**

- Developed and published privacy threat framework focused on web/IoT/mobile privacy control interfaces.

Teaching Assistant – **Jan. 2024 – Sept 2024**

- **Summer 2024** - Math & CS for Machine Learning (ML Dept.), **Fall 2024** - Privacy in the Digital Age (Public Policy & Information Systems Dept.), **Spring 2025**: - Engineering Privacy in Software (CS Dept.)

UPenn Computational Social Science Lab **Philadelphia, PA**
Research Data Scientist, Data Engineer **April 2022 – Aug 2023**

- Developed real time ETL pipeline for >50 TB of unstructured TV transcripts using Apache Spark, resulting in over 50% monthly savings in AWS EC2 compute and S3 storage costs.
- Developed Selenium web-crawler to scrape YouTube videos - scaled to review over 13K hours of content.
- Evaluated interpretability of hierarchical clustering of online political content using NLP-based embeddings.
- Managed compliance, privacy, data governance and access policies for AWS hosted data warehouse.
- Published research on online political radicalization at leading social science conference. **Publication here**.

Point72 Asset Management **New York City, NY**
Portfolio Data Scientist **Sept 2019 - April 2022**

- Built mixed-effect regression models to forecast monthly crude oil production across the Texas Permian Basin.
- Developed Random Forest models to forecast same-store sales and other KPIs to inform investment decisions.
- Delivered real-time analytics with event-driven ETL pipelines using Airflow, Jenkins, Databricks, and Tableau.

AllianceBernstein Asset Management **New York City, NY**
Investment Research Associate, Fixed Income **June 2017 - August 2019**

- Evaluated reliability of bond pricing mixed-effect regression model within niche market segments.
- Developed real-time corporate bond supply and demand metrics from unstructured order book database.

PROJECTS

Fairness in Federated Deep Learning (CMU) **Sept 2024 - Dec 2024**

- Developed a novel federated learning approach for training neural networks using PyTorch and Flower AI - balancing model accuracy with group fairness measures in crime recidivism prediction context.

Location Privacy for Vehicle-to-Vehicle Communication (CMU) **Aug 2024 - Dec 2024**

- Developed a differential privacy mechanism for vehicle-to-vehicle communication to mitigate the risk of re-identification attacks - tested privacy and utility tradeoff on simulated vehicular network data.

SKILLS

Scripting: Python, R, Scala, Bash **DevOps:** Git, Airflow, Docker, Jenkins, SQL **Cloud:** AWS (S3, EC2, Lambda, SNS, SES, Glue, Athena, IAM) **Database:** PostgreSQL, MongoDB **Machine Learning:** Sklearn **Deep Learning:** PyTorch. **Web Development:** Flask, Plotly/Dash **Web Scraping:** Selenium