

DREW LAIRD

drewlaird01@gmail.com | <https://drew-laird.github.io> | (860) 597-7778 | Washington, DC 20009

SKILLS AND INTERESTS

Programming Languages: Python, SQL, R

Machine Learning: Pandas, Scikit-learn, Transformers, NumPy, TensorFlow, PyTorch, NLTK, spaCy

Miscellaneous: AWS, Matplotlib, Power BI, Git, ArcGIS Enterprise

Interests: Cooking, Spanish, Competitive Cycling, Film Photography, Surfing, Bass Guitar

WORK EXPERIENCE

Snorkel AI, Washington, DC | *AI Engineer*

July 2025 – Present

- o Created an evaluation framework for Terminal-Bench environments to scalably produce programming solutions to human-generated finance problems for reinforcement learning applications. Validated the difficulty of the problems relative to several reference models, and error modes during the solution trace.
- o Develop synthetic data generation pipelines for frontier language modeling clients. Responsible for designing agentic system architecture, writing blueprints and rubrics to guide LLM generation, and creating evaluation frameworks to validate data quality.
- o Built and tested custom LLM-as-a-judge evaluators and deployed in production to guide both human-generated data pipelines and synthetic data generation pipelines. Designed a custom evaluator around a fine-tuned LLM to detect the presence of AI-generated content in human-produced data.
- o Generate synthetic data for an agentic system that provides daily planning and scheduling guidance. Utilized deep research capabilities and chained LLM calls to design proprietary data points for plan content in highly structured JSON format.

MITRE, McLean, VA | *Machine Learning Engineer*

January 2024 – July 2025

- **Modeling Minds - Using LLMs to Decode Media Influence on Human Behavior**
 - o Awarded \$100k to develop an experimental framework for developing AI that serve as human proxies for understanding behavioral change in response to media consumption and propaganda.
 - o Built a dataset of Latin American news for pre-training and a fine-tuning dataset of human survey data to create agents aligned to the language and cultural nuances of different Latin American countries.
- **Environmental Intelligence Platform**
 - o Served as the lead data scientist and technical point of contact on a real-time natural disaster risk analysis platform. Presented to government sponsors including the Department of Homeland Security Science and Technology Directorate, the Cybersecurity and Infrastructure Security Agency, and the Department of Defense.
 - o Gathered open-source historical sensor, satellite, and dam specification data around 300 known incidents. Expanded the open-source training pool using synthetic data generation and produced a classification model capable of predicting overtopping incidents with over 90% accuracy, recall, and precision.
 - o Geospatially visualized model predictions for 3000+ dams in North and South Carolina in near real-time using in-house ArcGIS Enterprise infrastructure and live satellite and weather predictions.

EDUCATION

The College of William & Mary, Williamsburg, VA

August 2019 – May 2023

- Bachelor of Science in **Computational and Applied Mathematics and Statistics (CAMS)**
 - o **Cumulative GPA:** 3.79/4.0
- **Awards:** Dean's List Fall 2020, Spring 2021, Fall 2021, Spring 2023; Rocket Pitch Winner; Employee of the Semester Spring 2021; Lead Role in Filipino Play; featured in school newspaper, podcasts, and the Hartford Courant for cycling.
- **Extracurriculars:** Captain of W&M Cycling Team; Student Leader of Bike Alliance; Area Supervisor at SUE

PROJECTS

Quantifying Cycling Training Distribution and Efficacy

March 2023 – August 2023

- Produced two research papers about quantifying and evaluating eighty-four hours of self-recorded competitive cycling training data. Cleaned and feature-engineered activity files collected on a variety of biometric and physical sensors.
- Proposed techniques to identify optimal intensity distributions, and model cardiovascular fitness on a day-to-day basis.