Samantha Evans, PhD

Data Scientist

(732) 718-3793 ◆ samanthagevans7@gmail.com ◆ Parlin, NJ 08859
LinkedIn ◆ Portfolio ◆ GitHub

Professional Summary

Data scientist and applied mathematician with expertise in machine learning, statistical modeling, and simulation-based analytics. Skilled in building and debugging end-to-end ML pipelines, developing predictive models, and extracting insights from large and complex datasets. Experienced in high performance computing, scientific computing and data visualization. Collaborative problem solver who bridges technical depth with clear communication to diverse stakeholders. Recent contributor to LLM benchmarking at Snorkel AI and developer of multi-language ML workflows for advanced simulation data.

Areas of Expertise

- Python
- Machine Learning
- Predictive Modeling
- Statistical Modeling
- Data Analytics
- Pytorch
- ML Pipelines
- Deep Learning
- matplotlib
- pandas
- NumPy
- scikit-learn
- Data Visualization
- Data Mining
- Experimental Design
- ETL
- Numerical Methods
- HPC
- MATLAB
- R

Experience

NHL Goal Probability Modeling - *Personal Project*

GitHub Repository | September 2025 - Present

- Conducted exploratory data analysis on NHL shot data from MoneyPuck.com to uncover patterns and scoring trends.
- Built logistic regression models to predict goal probabilities, refining models through feature engineering and evaluation.
- Developing nonlinear approaches (in progress) to improve prediction accuracy and capture complex gameplay dynamics.
- Produced visualizations and documentation to communicate results and ensure reproducibility.

Expert Contributor Mathematics - *Snorkel AI*

Remote (US) | January 2025 - Present

- Authored and validated original, graduate-level problems to test and enhance LLM fluency and reasoning capabilities.
- Contributed to a proprietary dataset of high-rigor challenges used in AI fine-tuning, benchmarking, and reinforcement learning.
- Uncovered failure modes in deductive logic, computation, and mathematical reasoning.
- Advanced the evolution of AI's mathematical cognition, influencing model fine-tuning strategies and shaping the next generation of AI reasoning benchmarks.

Research Assistant - *Department of Mathematics NJIT*

Newark, NJ | September 2019 - August 2024

- Built and maintained end-to-end ML pipelines for training neural networks on simulation data across Python and MATLAB.
- Designed and implemented a custom solver for two-phase flow, improving computational efficiency via boundary integral methods.
- Optimized simulation code for HPC clusters, significantly reducing runtime and scaling performance for large datasets.
- Developed a high-accuracy numerical method for simulating electroconvective flows using spectral representations.
- Produced visualizations and simulation videos to support research insights and cross-team collaboration.
- Collaborated across disciplines, translating mathematical models and computational results into actionable insights.
- Led recitations and coding labs for undergraduate courses (20-40 students per semester); created assignments, debugged student code, and clarified complex methods.

Education

Ph.D. - Applied Mathematics | New Jersey Institute of Technology, Newark, NJ | August 2024
B.S. - Mathematics (Minor: Statistics) | University of Rhode Island, Kingston, RI | May 2019

Publications

S. G. Evans, M. Siegel, J. Tausch, M. R. Booty. "A fast mesh-free boundary integral method for two-phase flow with soluble surfactant." Manuscript submitted for publication, Journal of Computational Physics and arXiv: https://arxiv.org/abs/2506.11282, 2025.