

Raymundo Navarrete

Data Scientist & Machine Learning Engineer

raymundo.navarrete.28@gmail.com • (+1) 614-842-3616 • [linkedin/raymundo-navarrete](https://www.linkedin.com/in/raymundo-navarrete)

SUMMARY

Skilled machine learning researcher/engineer with 10+ years of experience designing and implementing ML models, and 5+ years of experience bringing ML solutions to production. Proficient in handling and analyzing large data sets and cloud/distributed training. Comfortable working in high-performing teams. Experienced in risk assessment and modeling within the FinTech industry and government contracting. US Citizenship.

WORK EXPERIENCE

Senior Data Scientist • [Best Egg \(Personal Loans\)](#) – remote • 03/24 - present

- Led end-to-end platform migration and automation of **loan loss forecasting** processes, delivering more reliable dashboards and reducing team workload by 2 days per month.
- Designed and implemented an **alternative-data second-look model** for a flexible rent product, increasing application approvals while maintaining stable default rates.
- Developed next-generation **early delinquency model** and streamlined model development workflows for the flexible rent product, accelerating the full model deployment pipeline.

Skills: Snowflake, AWS Metaflow/Outerbounds, Mode Analytics (dashboard), LLMs

Senior Data Scientist • [Battelle \(Science & Technology\)](#) – Columbus, OH • 03/23 - 02/24

- Led the procurement of **\$1 million in funding** and managed a cross-functional team to develop a **predictive risk assessment model** for the **Environmental Protection Agency**, delivered within budget and on time.
- Spearheaded the development of **3 risk-assessment models** for diverse contracts, employing **advanced modeling techniques** to ensure client satisfaction.
- Leveraged **prompt engineering, memory management, and fine-tuning large language models** to identify and assess national security risks associated with chemical synthesis methods for explosive production by adversaries.
- Skills: C++, Excel/VBA, setup databases, GIS, Azure, LLM prompting & tuning

Machine Learning Researcher • [Upstart \(Personal Loans\)](#) – Columbus, OH • 04/21 - 02/23

- Developed **state-of-the-art machine learning techniques** to improve the company's flagship product, **increasing quarterly performance gains by 100%** and allowing team to reach quarterly goals.
- Conducted time-sensitive data analysis to identify and rectify under-performance of **personal-loan underwriting model** model during periods of economic turbulence.
- **Performed data cleaning and analysis to identify patterns of fraudulent activity.**
- Skills: Python, SQL, scikit-learn/xgboost, TensorFlow/Keras, AWS S3/EC2, airflow

Postdoctoral Researcher • [University of Arizona](#) – Tucson, AZ • 08/18 - 03/21

- Published multiple **machine learning publications** in top journals.
- Skills: Python, recursive/graph neural networks, ML visualization

EDUCATION

PhD in Applied Mathematics • [University of Michigan](#) • 08/13 – 08/18

- Received award for best applied mathematics thesis ([link](#)).

PROJECTS/SKILLS/PUBLICATIONS

Machine Learning

- Designing and deploying large language models
- Model development (pre-processing, data engineering, prototyping, feature selection, modeling, hyper-parameter tuning, explainability and governance).
- Supervised/unsupervised/reinforcement learning.
- Classification, regression, time series prediction, survival analysis, and anomaly detection using xgboost, deep/recursive/graph neural networks, and support vector machines.
- Graph/network visualization and dimensionality reduction (ML visualization), word embeddings (NLP), optical character recognition (image recognition).

Programming

- **Programming languages:** Python/Julia, C/C++/C#, Excel/VBA, HTML/CSS/Java
- **MLOps:** AWS S3/EC2/Batch/Lambda, Microsoft Azure,
- **Databases:** SQL, PostgreSQL, Redshift, SQLAlchemy, Snowflake
- **Cloud computing:** MLFlow, Airflow, Metaflow, Outerbounds
- **Distributed computing:** OpenMP, MPI, Dask
- **Machine learning:** scikit-learn, xgboost, TensorFlow, Keras, networkx
- **GitOps:** Git, github, Docker, Portainer, Kubernetes
- **LLMs:** openai, flask, autogen, llm, gradio, open gpt, context injection, fine-tuning
- **Dashboards:** PowerBI, Mode Analytics, plotly/dash

Publications

- Miller, J., Huroyan, V., Navarrete, R., Hossain, I., & Kobourov, S. (2024). ENS-t-SNE: Embedding Neighborhoods Simultaneously t-SNE. In *17th IEEE Pacific Visualization Symposium (PACIFICVIS)*. IEEE.
- Hossain, M. I., Huroyan, V., Kobourov, S., & Navarrete, R. (2020). Multi-Perspective, Simultaneous Embedding. *IEEE Transactions on Visualization and Computer Graphics*.
- Navarrete, R., & Viswanath, D. (2020). Prevalence of delay embeddings with a fixed observation function. *Physica D: Nonlinear Phenomena*, 414, 132697.
- Navarrete, R., & Viswanath, D. (2019). Delay embedding of periodic orbits using a fixed observation function. *Physica D: Nonlinear Phenomena*, 388, 1-9.
- Navarrete, R., & Viswanath, D. (2018). Prediction of dynamical time series using kernel based regression and smooth splines. *Electronic Journal of Statistics*, 12(2), 2217
- Navarrete, R., & Viswanath, D. (2016). Accuracy and stability of inversion of power series. *IMA Journal of Numerical Analysis*, 36(1), 421-436.