Rami Abou-Shamalah

(613) - 261 - 3542 | Ramiaboushamalah@gmail.com | [Website](#) | [Github](#) | [LinkedIn](#)

Education

BSc, Applied Mathematics   2020
Western University

MS, Applied Data Science   Expected 2026
University of Michigan

Experience

Data Scientist   June 2023 - Present
Sander Geophysics   Ottawa, ON

- Designed and implemented a data pipeline for time series anomaly detection, surpassing human accuracy
- Created Python-based data extraction scripts that convert archived data into structured files, then orchestrated a multi-stage ETL (extract, transform, load) process using Pandas and NumPy.
- Deployed a neural network model to production after exhaustive testing and edge-case validation.
- Saved $10,000 annually by optimizing batch processing, data cleaning, and model deployment workflows.

Lead Data Analyst   December 2021 - June 2023
Healthcare Systems R & A   Remote in Canada

- Led a team of 5 to collect, store, and implement QA/QC protocols for complex datasets, ensuring high-quality inputs for ML models.
- Migrated datasets from spreadsheets to NoSQL databases on AWS, improving scalability and accessibility.
- Built efficient Python pipelines for data cleaning, boosting processing speed by 30%.
- Delivered AI/ML insights to investors, securing new funding and advancing strategic initiatives.

Skills

- **Languages and Libraries:** Python, SQL, Batch Scripting, HTML, CSS, JavaScript
- **Libraries:** NumPy, Scikit-learn, Keras, Pandas, Matplotlib, Seaborn
- **Cloud & MLOps**: AWS (S3, ECS, ECR, EC2, SageMaker, Lambda) Docker, CI/CD, Git, API Development

Certificates

- AWS Machine Learning Speciality (ML-C01)
- Deep Learning & Machine Learning Specialization (Coursera)
- Front End Development & JavaScript (FreeCodeCamp)

Projects

Outperforming Humans in Anomaly Classification of Geophysical Data Using Machine Learning

- Developed on-premise ML pipeline using TensorFlow, achieving 95% accuracy in anomaly detection
- Feature engineering the data using signal processing methods such as advanced time series decomposition (seasonality, trend, residual) and moving window statistical values (mean, std, max)
- Presented the results at a KEGS (Society of Exploration Geophysicists) in September 2024

AWS MLOps Pipeline for Geological Prospecting

- Building a live pipeline pulling data from Sigeom (public geoscience data) into AWS S3, automating cleaning and preprocessing via Dockerized jobs.
- Orchestrates model training, validation, and deployment, providing region-specific gold prospectivity predictions across Quebec.
- Implements CI/CD, manual approval checkpoints, and API endpoints for real-time predictions