

MICHAEL SHEROUBI

devsheroubi@gmail.com | msheroubi.github.io | [GitHub/LinkedIn](#)

EDUCATION

Bachelor of Science in Computer Science

University of British Columbia

- Honours Program | Minor in Mathematics / Data Science* | 3.90 GPA on 4.33 scale

Certifications: CompTIA Security+, Neo4j Certified Professional, Google Data Analytics

PROFESSIONAL EXPERIENCE

Machine Learning Engineer. *EdgeTunePower*

2024

- Designed, built and optimized deep reinforcement learning (DRL) models for real-time optimal power flow on a simulated electric grid.
- Developed an anomaly detection system using an Autoencoder with LSTM regressor to flag anomalies from a real-time simulated electric grid.
- Established ETL pipeline connecting data from our simulator to the Azure cloud through Event Hub and Kafka.

Data Analyst. *Asurion*

2023

- Fine-tuned a large language model (LLM) to preemptively detect prompting ineligible products.
- Developed a clustering algorithm using LLM tokenization and cosine similarity, reducing the workload of the Digital Mapping team by an average of 88% per batch.
- Designed a suite of testing scripts to help automate QA by cross referencing data across multiple batches.
- Contributed to a collaborative team project focused on developing a machine learning classifier.

Data Analyst & Support. *Clearmind International*

2020 - 2023

- Developed an ETL solution using Python to migrate over 7 years of historical data; from legacy systems & csv files to a cloud-based solution.
- Designed OLTP database to handle simultaneous transactions by 20+ instructors while maintaining ACID compliance & data integrity.
- Normalized transcript data to improve query performance, reducing query run-time by over 75% and increasing accuracy of search results.
- Optimized registration process by modeling flow of data & trimming waste, cutting operational costs by 23%.
- Ensured enterprise-wide compliance with federal & international data protection regulations.

Data Engineer. *Trellis*

2019 - 2020

Warehouse Engineer / Team Lead

- Designed, developed, and implemented a data warehouse solution using AWS Redshift, combining data from 5 disparate sources to prepare them for analysis.
- Developed an ETL solution to migrate data in batches from mongo DB to the data warehouse using Python scripts & task scheduler.
- Created and executed automated test plans, cases, and scripts to validate data integrity, and identify software issues & their root causes.
- Designed and optimized complex SQL procedures called by the company's internal dashboard.

SKILLS

Programming Languages: Python, JavaScript/TypeScript, Java, SQL, Cypher

Platforms & Services: AWS SageMaker, IoT SiteWise, Redshift | Azure ML, Event Hubs, Stream Analytics | MongoDB

Libraries & Tools: PyTorch, OpenCV, Langchain, Scikit-learn, Scrappy, Pandas, Postgres, AutoGen, Gym

Theory: Deep RL, Policy gradient, IoT, Dimensionality Reduction, Data Analysis, Anomaly Detection

Languages: English (Fluent), Arabic (Fluent), French (Intermediate)

