

## SKILLS

- **Core (hands-on):** Python, Pandas, NumPy, SciPy, Jupyter, scikit-learn, Model evaluation & optimization, Anomaly detection, Statistical modeling, t-tests, ML models, LLM fine-tuning (LoRA/QLoRA), Git, Agile/Scrum
  - **Working knowledge:** SQL, PySpark, PyTorch, Matplotlib, Seaborn, D3.js, Docker, Java, Go, Azure ML,
  - **Exposure:** BigQuery/GCP, Kubernetes
- 

## PROFESSIONAL EXPERIENCE

### **Empower Analytics - Worcester, U.S. (Part-time)**

**01/2025 – 03/2025**

Machine-learning analytics firm providing pricing optimization for transportation and logistics clients

#### **Junior Data Analyst (ML Performance & Reporting)**

- Stabilized a live performance dashboard for an ML pricing tool by migrating reporting logic from Google Sheets to Python (Spark/Pandas) and adding data validation to handle inconsistent client inputs.
- Analyzed edge cases and data quality issues to improve feature reliability and reduce dashboard breakages during live updates.
- Produced weekly model performance reports with visualizations and t-tests for regular client updates.

**Tech Stack:** Python, PySpark, Pandas, SQL, Matplotlib, D3.js, Google Sheets API

### **Hanover Insurance Group - Worcester, U.S. (Internship)**

**05/2022 – 08/2022**

Major property & casualty insurer providing personal and commercial coverage nationwide

#### **Software Developer Intern**

- Built a Python/Pandas automation to parse Excel-based insurance claim data and convert it into an internal-system-ready format, significantly reducing manual entry time ( $\approx 15x$  faster vs manual processing).
- Supported data preparation for a reporting workflow used in the team's BI pipeline (Power BI exposure).
- Worked within an Agile sub-team of 5 (mentor + PM + intern pair), delivering increments and integrating scripts into internal processing flow.

**Tech Stack:** Python, Pandas, Excel, Power BI (basic)

---

## SELECTED PROJECTS (Industry & Research)

**2023 – Present**

### **UT Austin Lab — Computer Vision ROI Identification (Ongoing, part-time)**

- Developing an ML-assisted workflow to identify regions of interest (ROIs) in microscopy brain-slice images, reducing reliance on manual expert annotation for astrocyte-related analysis.
- Designing a human-in-the-loop feedback loop (expert yes/no validation  $\rightarrow$  iterative retraining) with a simple upload/review flow to support continuous model improvement.

**Tech Stack:** Python, PyTorch, Image Processing, Experiment Tracking, Docker

### **PayPal — Fraud & Risk Analytics (NDA) | 2025**

- Built an end-to-end LLM fine-tuning workflow for credit risk assessment using LoRA/QLoRA, including data preparation and prompt design under strict confidentiality constraints.
- Benchmarked fine-tuned LLMs against baseline ML models (e.g., Random Forest, SVM, XGBoost) using F1, KS, and MCC, including multi-seed stability checks to validate robustness.

**Tech Stack:** Python, PyTorch, Transformers, LoRA/QLoRA, scikit-learn, Azure AI Foundry, HPC

### **Fastly — Network Traffic Anomaly Detection (Industry Project via University) | 2023**

- Implemented anomaly detection using KNN (Python) and Mean Absolute Deviation (MAD) to flag outliers in BGP network traffic data.
- Built a lightweight web interface using JavaScript + D3.js to visualize traffic spikes and highlight anomaly timestamps for faster investigation.

**Tech Stack:** Python, scikit-learn, JavaScript, D3.js, BGP parsing, Go

---

## EDUCATION

### **Worcester Polytechnic Institute**

Master's Degree in Artificial Intelligence (in progress, expected May 2026)

Bachelor's Degree in Mathematics (With Honors)

Bachelor's Degree in Computer Science (With Honors)

---

**2024**

**2024**