

Dhruv Shah

437-256-1066 | dhruvshah38@gmail.com | [linkedin.com/in/dhruvshah38](https://www.linkedin.com/in/dhruvshah38) | github.com/dhruv-shah

EDUCATION

University of Waterloo

Bachelor of Computer Science

Waterloo, ON

Expected Sep 2027

Wilfrid Laurier University

Bachelor of Business Administration

Waterloo, ON

Expected Sep 2027

EXPERIENCE

Data Analyst – Automation

Sep 2025 – Feb 2026

Independent Electricity System Operator (IESO)

Toronto, ON

- Analyzed **10+ years** of operational data across **6 hydroelectric generators** and flagged payment exceptions affecting **\$100K** by building a Python and Excel Power Query pipeline to compare performance against contractual thresholds.
- Saved **8 hours per month** by building documented automation tools that archived team notes and surfaced upcoming deliverables across configurable timeframes.

Software Engineer

Jan–Apr 2024; Jan–Apr 2025

KORE Solutions

Toronto, ON

- Reduced algorithm runtime by **97%**, **from five minutes to ten seconds**, by replacing an index-based collection with a write-optimized hash map.
- Restored commission accuracy across **2,000+ zero-cost records** by executing targeted fixes in a GemStone production database, eliminating year-end accounting mismatches.
- Reduced end-to-end test failures by **80%** and CI/CD runtime by **40 minutes** by developing Cypress test coverage for critical workflows.

Accounting Analyst – Automation & Data

Jan 2023 – Apr 2023

HTS

Toronto, ON

- Cut invoice-processing time by **50%** and manual-entry errors by **70%** by automating client invoice delivery with JavaScript and Selenium.
- Reduced ageing-report preparation from **2 hours to 5 minutes** by automating report generation and account visualizations with Python, pandas, and Matplotlib.

PROJECTS

Stem Cell Differentiation Pipeline | *R (Seurat), Python, Nextflow, Docker, Cell Ranger*

Jun 2026 – Present

- Built an automated pipeline that processed **50 GB of raw sequencing data** from approximately **120,000 cells in under two hours**.
- Implemented data-quality and normalization steps that captured the expected transition from **stem cells to heart-muscle cells** between day 0 and day 16 samples.

F1 Sponsor Stock Returns | *Python, pandas, statsmodels, yfinance, Matplotlib*

Jun 2026

- Built a reproducible Python event-study pipeline analyzing **778** sponsor-race observations by joining Formula 1 results with post-race stock returns from Yahoo Finance.
- Found race wins translated into a **sub-0.5% next-day sponsor-stock movement** and near-zero three-day reaction after controlling for sponsor and race fixed effects.

Gene Sequence Classifier | *Python, PyTorch, Biopython, NumPy, SQL*

Apr 2024

- Achieved **75% classification accuracy** across human, bacterial, viral, and fungal genomic sequences by training a Transformer model in PyTorch on a custom Biopython encoding pipeline.
- Designed a genomic encoding pipeline and SQL schema for efficient storage and retrieval of sequence data.

TECHNICAL SKILLS

Languages: Python, R, Bash, SQL, C++, C, JavaScript, Java

Bioinformatics: Seurat, Cell Ranger, Nextflow, Biopython

Tools: Git, Docker, Excel, Power Query

Libraries: PyTorch, pandas, NumPy, Matplotlib