Nihal Vellaramkallingal Sulaiman

Data Analyst · Business Intelligence · Machine Learning

Niagara Falls, ON · nihalunfc@gmail.com · Authorized to work in Canada · Available immediately for full-time roles

nihalsulaiman.com · github.com/nihalunfc · linkedin.com/in/vsnihal · kaggle.com/vsnihal

SUMMARY

Recent Master's graduate in Data Analytics (University of Niagara Falls Canada, October 2026) with an M.Sc. in Physics and a strong grounding in statistics, mathematics and modelling. I build end-to-end analytics solutions in Python, SQL, Power BI and Excel: star-schema data warehouses, automated data-quality checks, KPI dashboards, demand forecasts and optimisation models. Currently a remote Data Engineer at Shipd (DataCurve), creating training data for AI models after being recruited on the strength of my Kaggle results (top 10 to 15% finishes in global competitions). Holder of five professional certificates in data analytics and business intelligence from Google, SAS and Microsoft, and an experienced writer who explains data clearly to non-technical audiences.

EXPERIENCE

Data Engineer (remote)

Feb 2026 – Present

Shipd by DataCurve, San Francisco, CA, USA

- Recruited directly by the company after a top-15% Kaggle competition finish, to build high-quality training data for large language models and AI coding agents (Project Eris).
- Design complex programming and data challenges, with verified reference solutions, that expose reasoning gaps in advanced AI agents.
- Write, test and refine machine-learning code used to train and fine-tune models, meeting strict standards for correctness and quality.

Data Annotator (remote)

Jun 2025 – Sep 2026

TELUS Digital, Canada

- Labelled and quality-rated multimodal AI training data (text, image and video) on the Multimango and Parimango platforms.
- Performed entity tagging on social-media content, video interleaving, and image and video quality rating against detailed guidelines, keeping accuracy consistent at volume.

Content Writer (full-time, remote)

Apr 2019 – Apr 2024

BrightMedia International, Calicut, India

- Researched and wrote news articles, investigative reports, science-communication pieces and reviews, plus marketing copy, online ads and social-media content, to deadline for five years.
- Turned technical and scientific topics into clear, accurate writing for a general audience, the same skill I now use to present data insights to stakeholders.

DATA AND AI PROJECTS

Hospitality Data and Business Intelligence Platform (Projects HARLTON and H2)

Python, SQL, Excel, PuLP

- Built an end-to-end BI pipeline on a synthetic dataset of five hotels (about 640,000 bookings over four years): star-schema warehouse, ETL, ten automated data-quality checks and unit tests.
- Designed a Revenue Command Centre dashboard covering occupancy, ADR, RevPAR, booking pace against last year, cancellations, channel net rates and a back-tested demand forecast.
- Developing revenue and cost analyses (group displacement, rate recommendations, overbooking, energy costs) and a staff scheduler that turns forecast workload into shifts with mixed-integer optimisation, checks Ontario labour rules and exports to Excel.

Self-healing data pipelines (open source, MIT)

Python, multi-agent AI, Claude API, Dagster (planned)

- Designed a decentralised crew of nine AI agents with no central controller that detects, diagnoses, fixes, tests and re-runs failures in a daily ETL pipeline, with every step logged and reversible and production merges left to a person.
- Simulated a year of failures: median time to fix fell from 9.2 hours to 1.1 hours and people-hours on breakages from 534 to 121, for about $2 a year in model cost (simulated figures; a measured reference pipeline is next).

Kaggle competitions

Python, PyTorch, XGBoost, CatBoost, LightGBM

SKILLS

Analytics and BI

Power BI, Tableau, Excel (advanced), KPI design, dashboards, automated reporting, exploratory data analysis, statistical analysis (SAS)

Data and engineering

SQL (SQL Server / T-SQL, SQLite), ETL pipelines, star-schema data modelling, data cleaning, data-quality testing, Databricks, Google Cloud Platform, Git and GitHub

Programming

Python (pandas, NumPy, Polars), JavaScript, HTML, CSS, D3.js, XML, LaTeX

Machine learning and AI

Feature engineering, time-series forecasting, XGBoost, CatBoost, LightGBM, Optuna, PyTorch, reinforcement learning, LLM fine-tuning (LoRA), Gemini API

Optimisation and tools

Mixed-integer optimisation (PuLP), Microsoft Office, Google Workspace, AI coding assistants (Claude Code, Google Antigravity)

EDUCATION

Master's in Data Analytics

Apr 2025 – Oct 2026

University of Niagara Falls Canada, Ontario

M.Sc. Physics

2017 – 2019

Cochin University of Science and Technology, India

B.Sc. Physics (mathematics minor)

2013 – 2016

University of Calicut, India

CERTIFICATIONS

- Google Advanced Data Analytics Professional Certificate and Google Business Intelligence Professional Certificate (Coursera, 2026)
- SAS Statistical Business Analyst Professional Certificate (Coursera, 2026)
- Microsoft Data Visualization Fundamentals; Microsoft Data Preparation and Management (Coursera, 2026)
- Google Cloud Skill Badge: Prompt Design in Vertex AI
- Workplace: WHMIS with GHS, Worker Health and Safety Awareness, AODA Awareness, Delivery Driver Safety

CANADIAN WORK EXPERIENCE

Alongside my studies I have worked in Niagara's hospitality sector, building reliability, teamwork and

- Mallorn Challenge (top 15% of about 800 teams): detected rare astronomical events in noisy time-series data with under 1% positive cases, using physics-informed features and an XGBoost and CatBoost ensemble tuned with Optuna.
- Name That LLM (436th of 4,346 teams, top 10%, ongoing): fine-tuned Qwen2.5-7B and GLM-4-9B with LoRA on a 38,935-passage training set and built leak-free grouped cross-validation (0.663 balanced accuracy).
- NeurIPS 2026 Smart Buildings: designed a physics-informed HVAC control policy that beats the best public notebook (0.01244 vs 0.01853, lower is better), and found and reported a bug in the organisers' simulator.
- Also: Squares in Squares (top 13%), Enveda CASMI 2026 molecule identification, CAFA protein-function prediction (deep learning with extreme class imbalance), and Google MedGemma Impact Challenge.

WARL: Wildfire Autonomous Reinforcement Learning (GDG on Campus Build with AI Challenge 2026)

PyTorch, JavaScript, Leaflet.js, Gemini API

- Developed a decentralised drone-swarm simulation using a Deep Q-Network to route drones to wildfire choke points, building on my earlier multi-agent reinforcement learning model of ERCOT (Texas grid) energy data.
- Built a live web dashboard with satellite mapping and a Gemini-powered chatbot that explains the model's decisions in plain English.

customer service skills, plus first-hand knowledge of the hotel operations behind my BI projects.

Housekeeping Staff

Jun 2025 – Present

Canadian Niagara Hotels (Brock Hotel), Niagara Falls, ON

Waterpark Attendant

Oct 2025 – Jan 2026

Great Wolf Lodge, Niagara Falls, ON

Customer Service and Kitchen Staff

Aug 2025 – Oct 2025

Rapha Kerala Restaurant, Niagara Falls, ON

Kitchen Helper

Apr 2025 – Jun 2025

Mr Johns Kerala Kitchen, Brampton, ON

ACHIEVEMENTS AND LANGUAGES

- **Achievement:** Finalist, International Astronomy and Astrophysics Competition (IAAC) 2020
- **Community:** Member, Google Developer Group (GDG) on Campus, Niagara Falls (Oct 2025 – Present)
- **Languages:** English (fluent), Malayalam (native)