

Noel John

Data & ML Engineer

noeljohn19984501@gmail.com
linkedin.com/in/noeljohn1 · github.com/noeljo23
noeljohn.dev

Data and ML engineer who builds the data systems teams make decisions on, from pipelines over 10M+ records a day to the reporting and quality checks on top, and who coaches and leads along the way. 2nd worldwide at Ocean Hackathon 2025 and 1st place at the PowerTrust GenAI Hackathon.

EXPERIENCE

Data Analyst (Co-op), Global Operations · Geotab, Vancouver

Jan 2026 – Present

- Disproved the prevailing theory for a revenue projection overshoot by reconciling it against live data; the review owner made my monthly back-test a standing process.
- Explained a large recurring gap between two revenue systems, showing nearly all of it was a definition difference and pinning the real defects to specific code.
- Designed and built the data-quality layer and a Cloud Run failure-diagnosis service for the revenue pipelines, so problems surface before the monthly review.

Data Analyst / Business Analyst · Media.net, Mumbai

Aug 2021 – Aug 2024

- Automated the team's ETL workflows in Airflow and Alteryx, cutting manual workload by 40% and the monthly reporting cycle from six hours to under one.
- Cut the typical production data investigation from eight hours to two, and new-analyst onboarding time by 30% through the documentation I kept.
- Built Python and SQL pipelines over 10M+ records a day on BigQuery and Redshift, and lifetime value models for five apps with 5M+ monthly users.

Teaching Assistant · Northeastern University, Vancouver

Jan 2025 – Apr 2025

- Ran Python exercises, SQL tutorials and Tableau labs for 20 graduate students.

Global Recreation Ambassador · Northeastern University

Jan 2025 – Dec 2025

- Organised 20+ events for 150+ international students.

Instructor · Try Catch Classes, Mumbai

May 2023 – May 2024

- Trained working professionals in Alteryx, SQL and Tableau as a then-certified Alteryx instructor; 8 in 10 moved into analytics roles within three months.

PROJECTS & HACKATHONS

Ocean current classifier · Ocean Hackathon 2025, 1st in Canada and 2nd worldwide at the Brest Grand Finale

Oct – Dec 2025

- Built the vision model and data pipeline for an EfficientNet-B4 classifier that predicts whether tides, wind or river outflow drives hourly HF-radar current maps; 76.2% test accuracy.

Research assistant for solar investment · PowerTrust GenAI Hackathon, 1st place

Apr 2026

- Source-cited research assistant for solar investment in 10 emerging markets. Five agents check each fact against two official sources before indexing, and it refuses to answer without one.

Underwater 3D reconstruction benchmark · Research

Jun – Jul 2026

- Geometric benchmark of underwater 3D reconstruction on synthetic scenes with exact ground truth, 210 generic vs water-aware 3DGS runs over 3 seeds. Water-aware 3DGS helped only on high-relief scenes in clear to moderately murky water.

Other hackathons

- Visualization Hackathon, Northeastern (Dec 2024): led a team of 3 to 2nd of 12 teams, analysing 50,000+ sales records.
- InnoFest 2025: built the video path for a WorkSafeBC fall-protection checker, using Azure Functions and Video Indexer.
- Northeastern Global Innovation Challenge, Dentons (Mar 2026): a deal handoff tool for an international law firm, on FastAPI.

EDUCATION

MSc, Data Analytics Engineering · Northeastern University, Vancouver · GPA 3.85

Sep 2024 – Apr 2027

B.E., Electronics Engineering · University of Mumbai

Aug 2017 – Jun 2021

SKILLS

Languages and infra · Python, SQL, R, Docker, Kubernetes, Linux, Git, GitLab CI

Data · BigQuery, Redshift, PostgreSQL, Airflow, Alteryx, data modelling, data-quality testing

Cloud · Google Cloud Run, Cloud Composer, AWS (S3, Lambda), Azure Functions, SLURM on H200 GPUs

ML and LLMs · scikit-learn, PyTorch, TensorFlow, XGBoost, MLflow, SHAP, A/B testing, forecasting, RAG, LangGraph

BI and analytics · Tableau, Power BI, Looker, Superset, Mixpanel, Amplitude, Google Analytics, Streamlit