

Tan Dai Ngo

📞 250-661-9530 | 📍 Victoria, BC | ✉ ngotandai95@gmail.com | 🔗 linkedin.com/in/ntdai95/ | 🌐 ntdai95.github.io

Technical Skills

- **Programming & Databases:** Python, SQL, R, MATLAB, Cypher, Java, MySQL, MongoDB, Cassandra, Neo4j
- **Frameworks & Tools:** Power BI, Excel, Jupyter, Pandas, Numpy, SciPy, StatsModel, Scikit-learn, TensorFlow, Keras, LightGBM, XGBoost, CatBoost, Plotly, Matplotlib, Seaborn, AWS (EC2, S3, RDS), Docker, Jenkins

Education

Master of Engineering in Applied Data Science | University of Victoria **Sep 2025 - Present**

Master of Science in Computer Science | University of Chicago **Sep 2020 - Mar 2022**

Bachelor of Science in Economics | University of Washington **Sep 2015 - Aug 2019**

Work Experience

Data Analyst | Linh Dan Hotel *Ho Chi Minh City, Vietnam* | **Apr 2023 - Aug 2025**

- Built SQL ETL pipelines (20+ daily queries) and designed ERD-modeled schemas to integrate and clean 1,000+ operational and financial records into analysis-ready datasets.
- Automated reporting workflows with Python and Power BI, reducing manual reporting time by 30% and enabling structured monthly/quarterly reviews.
- Performed cost-benefit analyses and built Python analytics dashboards (Matplotlib, Pandas) to guide pricing strategy and optimize operational efficiency.

Computer Science Teacher | APU American International School *Da Nang City, Vietnam* | **Aug 2024 - Dec 2024**

- Instructed 50+ high school and college students in python programming and web development via practical lessons.
- Designed coding projects that apply real-world software engineering principles such as top-down design and OOP.
- Mentored students in debugging and code optimization, fostering analytical reasoning and critical thinking skills.

Software Developer | T-Mobile (BeaconFire Inc.) *Bellevue, WA, USA* | **Jun 2022 - Feb 2023**

- Collaborated with a team of 5 to develop enterprise microservices for the Roaming Business System (Spring Boot, Java Stream API), integrating with Jenkins CI/CD for deployments and Splunk for log monitoring.
- Designed and optimized Cassandra database schemas for partner and workflow management microservices, handling 100k+ records while maintaining query execution time under 3 seconds.
- Automated weekly reprocessing of 50+ roaming service tests with Kafka pipelines, cutting manual effort by 80%.

Projects

Predictive Modeling for Sustainable Fuel Blending **Shell.ai Hackathon 2025**

- Identified key component-property relationships through exploratory data analysis of a 65-feature fuel-blend dataset, enabling targeted feature engineering for predictive modeling.
- Engineered weighted property-based aggregates, cross-component interactions, and log-transformed features to model non-linear blend behaviors across 10 sustainable aviation fuel properties.
- Developed a stacked Ridge regression ensemble of gradient-boosted decision tree models trained with CatBoost, LightGBM, and XGBoost frameworks, optimized via Optuna hyperparameter tuning, achieving MAPE < 2.58%.

Employee Attrition Prediction **Google Advanced Data Analytics Capstone Project**

- Analyzed a 15,000-row HR dataset to identify drivers of employee attrition, uncovering that low satisfaction, overwork, and tenure patterns strongly correlated with departures.
- Engineered and evaluated predictive models (logistic regression, decision trees, random forests), achieving 96% test accuracy and 0.94 AUC in predicting attrition risk.
- Developed dashboards and feature plots to translate model outputs into actionable insights for strategic decisions.