

ORANICH JAMKACHORNKIAT

(332) 283-4130 | oj2191@columbia.edu | linkedin.com/in/oranich-jamkachornkiat | github.com/nniishhh

EDUCATION

Columbia University

MS, Operations Research

New York, NY

Expected Dec 2026

GPA: 3.72 | Coursework: Applied Machine Learning, Optimization Models & Methods, Stochastic Models, Analytics on the Cloud

Chulalongkorn University (Ranked #1 in Thailand)

BEng, Robotics & AI Engineering

Bangkok, TH

Oct 2023

GPA: 3.61 | Honors: First Class Honors, Academic Excellence Scholarship (2022), RoboCup - 2nd Place

EXPERIENCE

MTA - New York City Transit (Dept. of Subways, Performance Analysis)

New York, NY

Performance Analytics Intern

Jun 2026 - Aug 2026

- Linked 6M train-arrival records to 490K track-work orders in Python and SQL, then compared flagged runs with matched non-flagged runs by direction and time of day; isolated 63.7 seconds of added runtime per segment, roughly 70% longer than normal.
- Analyzed added runtime by department, location, and time of day, showing the impact was concentrated in one department (89 midday hours, 3.1 minutes per affected trip); leadership agreed to shift midday track work overnight.

SWAT Mobility (B2B Transport-Tech Startup, Enterprise Logistics Clients)

Singapore, SG

Data Scientist

Nov 2023 - Dec 2025

- Owned the full model lifecycle for a production DHL Turkey address service, fine-tuning BERT and RoBERTa for named-entity recognition to reach 98% extraction accuracy and hardening releases with CI/CD, unit tests, and Git workflows.
- Created a route-optimization model for 10K daily 7-Eleven deliveries with agglomerative clustering under cost, service-level, and capacity constraints, evaluating variants offline; cut planning time by 70% and operational cost by 8%.
- Visualized 2M GPS trajectory records in interactive geospatial dashboards (Kepler.gl, MovingPandas), revealing mobility patterns that helped Japanese transit operators cut inefficient off-peak trips by 30%.
- Automated planning KPI and data-validation monitoring with Python, Windmill, and Slack alerts, eliminating 4 hours of daily manual investigation and speeding cross-team issue triage.

Shopee (Southeast Asia's Largest E-Commerce Platform, Sea Group)

Bangkok, TH

Data Analytics Intern (Automation)

May 2022 - Aug 2022

- Benchmarked OCR vendors across Google, AWS, and Azure and added edge-detection and skew-correction preprocessing for variable-quality ID and bank documents, raising verification throughput by 85%.

Agoda (Online Travel & Booking Platform, Booking Holdings)

Bangkok, TH

Data Analytics Intern (Supply)

Jun 2021 - Nov 2021

- Implemented room-type mismatch detection over unstructured listing text with lemmatization, TF-IDF, and Levenshtein distance, validating against existing rules and improving detection by 120%.

PROJECTS

Airbnb Data Analyst Agent - Agentic AI, Columbia

[Project Link](#)

- Architected a multi-agent pipeline with the OpenAI Agents SDK that answers natural-language questions over 37K NYC listings and 985K reviews by generating and executing DuckDB SQL and Python, then returning charts and narrative insights through a React front end.

Credit-Risk Modeling for Kalamata Capital Group - Columbia

2025

- Trained XGBoost, CatBoost, and regularized logistic-regression models on 100K transactions with 200 engineered features, using TimeSeriesSplit cross-validation to prevent leakage and SHAP for feature attribution, achieving 0.82 AUC.

Facility Location & Capacity Expansion - Optimization Models & Methods, Columbia

2025

- Formulated two mixed-integer linear programs to site facilities and size expansions across New York State ZIP codes under piecewise costs, expansion caps, and separation constraints, yielding a \$322M minimum public expenditure.

SKILLS

Machine Learning: XGBoost, CatBoost, logistic regression, Random Forest, feature engineering, cross-validation, SHAP interpretability, transformer fine-tuning (BERT/RoBERTa), NER, NLP, TF-IDF.

Optimization & Statistics: Vehicle routing & scheduling optimization (VRPTW), mixed-integer linear programming, clustering, regression, statistical inference, matched-baseline comparison.

Programming & Data: Python (Pandas, NumPy, Scikit-learn), SQL, PySpark, Spark SQL, Trino, Airflow, Git.

Production: Docker, Google Cloud Run, Cloud Build CI/CD, Windmill, Slack alerting.