

Jonathan Lai

laijonathankl@gmail.com | (412) 450-7676 | [linkedin.com/in/jonathanlai](https://www.linkedin.com/in/jonathanlai)

EDUCATION

Carnegie Mellon University (CMU)

Pittsburgh, PA

BSc. Statistics & Machine Learning, Double Major in Information Systems

Expected 2026

- Relevant Courses: Introduction to Machine Learning, Database Design & Development, Advanced Methods for Data Analysis, Modern Regression, Foundations of Software Engineering

RELEVANT EXPERIENCE

Hong Kong Applied Science & Technology Research Institute (ASTRI)

Hong Kong

ML Engineering Intern

Jun 2024 – Aug 2024

- Analyzed network traffic big data for 5G network resource allocation optimization via RAN deployment
- Productionized ASTRI software architecture by integrating RAN, documenting references for transferability

OT&P Healthcare

Hong Kong

Data Consultant & Project Intern

Jun 2023 – Aug 2023

- Consulted operation team to identify strategic metrics, leveraging OT&P's database for dashboard prototype
- Built Tableau dashboards for 3 user groups to display KPIs like patient diagnosis, time allocation, & revenue
- Showcased to executives 9 BI dashboards that deliver actionable business-insights & optimize performance

Delight Labs

Hong Kong

Statistics Intern

Jun 2020 – Aug 2020

- Translated text data for RNN to classify from ASR transcript, improving user engagement across voice apps
- Launched Google Assistant app "English Football League" to interactively show teams' & players' statistics

Carnegie Mellon University, Department of Statistics & Data Science

Pittsburgh, PA

Teaching Assistant, Methods for Data Science & Statistics

Aug 2023 – Present

- Taught 40+ students with regression analysis & machine learning assignments on R during weekly lab
- Personalized feedback by facilitating course Q&A platform & grading homework, labs, & projects

PROJECTS - github.com/JonathanLai/PersonalProjects

NLP-Based Yelp Sentiment Prediction System

- Engineered sentiment analysis system for restaurant reviews by building a logistic regression classifier
- Harnessed feature engineering & stochastic gradient descent on Yelp's dataset in the NLP implementation

LLM with Adjustable Temperature

- Constructed a RNN model to perform sequential word predictions based on contextual embeddings
- Leveraged PyTorch for high-dimensional math & self-attention, which is integral to ChatGPT-like models

Discord App Database Design

- Architected relational & physical data model that mimics Discord app while normalizing for data integrity
- Simulated 10 sample user interactions by coding Python scripts run on PostgreSQL that satisfy user stories

Predicting Room Occupancy Report

- Evaluated best ML classification model for predicting room's occupancy using real data based on error rates
- Mitigated overfitting by refining classification models (e.g. LDA, QDA, classification trees, & logistic classifier)

1v1 Basketball Game

- Curated 6 animated & realistic NBA players to enhance authenticity for a versus AI basketball game
- Implemented AI opponent, physics, & basketball moves like dribbling, shooting & stealing using probability

SKILLS

Programming Languages: Python, R, SQL, C, HTML, CSS,

Tools: Tableau, MongoDB, PostgreSQL, PL/pgsql, NumPy, PyTorch, Git, Docker

Language: English (fluent), Cantonese (fluent), Mandarin (fluent)