

# Hongling Lei

honglinglei98@gmail.com | +1 (412) 805-8510 | [honglinglei.github.io](https://honglinglei.github.io) | [github.com/HonglingLei](https://github.com/HonglingLei)

## SUMMARY

Seasoned data scientist with expertise in Gen-AI application development using prompt engineering, LLM APIs, and object-oriented programming, plus strong skills in ML modeling and information retrieval. Seeking a Senior Associate role in AI/ML.

## PROFESSIONAL EXPERIENCE

### JPMorganChase

Associate (Data Scientist), Focused Analytics Solutions Team

New York, NY

Mar. 2023 – Ongoing

#### AI/ML:

- Developed and supported deploying Babelfish, a Gen-AI powered code translation and documentation tool that facilitates firm-wide codebase migration across languages and to the cloud, reducing manual effort by 95%
- Implemented batch-mode translation functionalities with automatic progress logging and error handling
- Delivered solutions for translating complex codebases with dependency graphs and Retrieval-Augmented Generation

#### Matching as a Service:

- Built a record linkage pipeline from the ground up, which leverages semantic similarity and vector retrieval to help Chase ingest diverse datasets through fuzzy matching, boosting search speed by 1000 times with over 97% accuracy

#### Auto Analytics:

- Independently created a Streamlit web app for dynamic, in-depth analyses of Chase customer digital journeys and engagement, guiding web design improvements to enhance customer retention and user experience

#### Consumer Bank:

- Developed Random Forest models to predict incremental deposit attrition rates following branch consolidations
- Proposed personalization strategies based on K-means customer segmentation, improving conversion rates by 5%
- Optimized banker efficiency by reducing meeting cancellations with data-driven solutions, increasing profit by \$60M

#### Leadership Initiatives:

- Led the Resourcing Committee's BAU workstream, managing rotation planning and staffing of 60+ team members
- Served as a team mentor in the Data for Good Hackathon, providing college students with guidance and support

### PPG Industries

AI/ML Intern, AI/ML Center of Excellence

Pittsburgh, PA

May 2022 – Aug. 2022

- Constructed 15 time-series forecasting models using ARIMA, Prophet, Exponential Smoothing and more to predict finished goods demand for inventory optimization, improving forecast accuracy by 17%
- Developed a Python-based Auto-Forecaster that takes any time series, experiments with 10+ forecasting algorithms, self-tunes hyper-parameters and recommends the best model, shortening iteration cycles from 2 weeks to 1 minute

### Tencent

Data Scientist Intern, Platform and Content Group

Shanghai, China

Apr. 2021 – July 2021

- Conducted causal inference with algorithms like Causal Bayesian Networks, X-Learner, and Causal Forest to identify key drivers of customer behavior on Tencent streaming platforms, increasing user satisfaction rating by 11%
- Deployed an inference pipeline that automates feature engineering, modeling, and experiment design to streamline A/B testing and optimize content recommendations, enhancing customer experience and engagement

## EDUCATION

### Carnegie Mellon University

Master of Information Systems Management. GPA 3.91 (Highest Distinction)

Pittsburgh, PA

Dec. 2022

- Coursework: Deep Learning, Computer Vision, NLP, Distributed Systems, Data Structures and Algorithms

### Xiamen University

Bachelor of Economics and Finance. GPA 3.81

Xiamen, China

June 2020

- Exchange Experience: University of California, Berkeley; Nanyang Technological University, Singapore

## SKILLS

- **Programming Languages:** Python, SQL, Java, R, Stata, HTML
- **AI/ML:** PyTorch, scikit-learn, pandas, sentence\_transformers, faiss, OpenCV, prompt engineering, Git, Bitbucket
- **Big Data/Cloud:** Spark, Snowflake, Hive, Teradata, Analyst Workbench, Google Colab, Azure