Hongling Lei

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

SUMMARY

Seasoned data scientist with expertise in Gen-AI application development using prompt engineering, LLM APIs, and OOP. Combined with strong skills in ML modeling and vectorized information retrieval, seeking a Senior Associate role in AI/ML.

PROFESSIONAL EXPERIENCE

JPMorgan Chase & Co.   New York, NY
Associate (Data Scientist), Focused Analytics Solutions Team   Mar. 2023 – Ongoing

- **AI/ML:** Developed and supported deploying Babelfish, a Gen-AI empowered code translation and documentation tool that facilitates firm-wide codebase migration across languages and to the cloud, reducing manual effort by 95%
- **Matching as a Service:** Built a record linkage pipeline from scratch that leverages semantic similarity (SBERT) and vector retrieval (FAISS) for fuzzy matching, boosting the search speed by 1000 times with over 97% accuracy
- **Auto:** 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:** Successfully delivered 3 projects, including attrition rate predictive modeling, banker efficiency optimization, and online shopping personalization strategies, achieving a $230M profit increase in total
- **Home Lending:** Improved customer segmentation and profiling using K-means clustering
- **Resourcing Committee:** Led the BAU workstream in rotation planning and staffing of 60+ team members

PPG Industries   Pittsburgh, PA
AI/ML Intern, AI/ML Center of Excellence   May 2022 – Aug. 2022

- Constructed 15 time-series forecasting models including ARIMA, Prophet, and Exponential Smoothing 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 min

Tencent   Shanghai, China
Data Scientist Intern, Platform and Content Group   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

DiDi   Beijing, China
Data Analyst Intern, Decision Support Team   Sept. 2020 – Nov. 2020

- Designed a market sizing model with SQL and Excel, accurately predicting driver and order growth rates during holidays and thus alleviating traffic burdens in 14 Chinese metropolitan areas
- Conducted large-scale analyses on business datasets using Apache Hive to advise strategic decision-making

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
Master of Information Systems Management. GPA 3.91 (Highest Distinction)   Dec. 2022

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

Xiamen University   Xiamen, China
Bachelor of Economics and Finance. GPA 3.81   June 2020

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

PROJECTS

- **Object Tracking System:** Target tracking for template matching and motion detection with the Lucas-Kanade method
- **Augmented Reality:** Image and video AR projection through interest point matching and homography estimation
- **Grocery Master:** Software that provides nearby availability of searched grocery products using real-time web scraping

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