Hongling Lei

Website | GitHub | LinkedIn | +1 (412) 805-8510 | hongling@andrew.cmu.edu

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
Master of Information Systems Management, Business Intelligence and Data Analytics. GPA 3.93 — *Aug. 2021 – Dec. 2022*
- Relevant Courses: Deep Learning, Computer Vision, Distributed Systems, Machine Learning, Unstructured Data Analytics (NLP), Object Oriented Programming in Java, Data Structures

Xiamen University — Xiamen, China
Bachelor of Economics, Finance. GPA 3.81 — *Sept. 2016 – Jun. 2020*
- **University of California, Berkeley**
  Semester Exchange, Statistics, Computer Science, Business. GPA 3.95
- **Nanyang Technological University, Singapore**
  Summer Exchange, Digital Marketing. GPA 5.0

PROFESSIONAL EXPERIENCE

PPG Industries — Pittsburgh, PA
AI/ML Intern, AI/ML Center of Excellence — *May 2022 – Aug. 2022*
- Constructed time-series forecasting models including ARIMA, Exponential Smoothing, and Prophet to predict finished goods demand for inventory optimization, improving the status-quo forecast accuracy by 17%
- Boosted productivity by developing a Python-based Auto-Forecaster that takes any time series, experiments with over 10 forecasting algorithms, self-tunes hyper-parameters, and recommends the best model
- Derived key drivers of demand fluctuations with explainable AI techniques like SHAP and LIME

Tencent — Shanghai, China
Data Scientist Intern, Public Data Science Department — *Apr. 2021 – Jul. 2021*
- Conducted causal inference with algorithms like Causal Bayesian Networks, X-Learner, and Causal Forest to analyze reasons behind consumer behavior on Tencent's streaming platforms, significantly improving user experience
- Deployed an inference pipeline that automates feature engineering, machine-learning modeling, and future interventions, putting everything into production for practical decision-making with live data

DiDi Global — Beijing, China
Data Analyst Intern, Decision Support Department — *Sept. 2020 – Nov. 2020*
- Supported business and strategy decisions by extracting and analyzing billions of data points on Apache Hive
- Designed a market sizing model with SQL and Excel, correctly predicting driver and order growth rates during holidays and thus alleviating traffic burdens in 14 Chinese metropolitan areas
- Created and monitored business dashboards that can update data and visualize product metrics weekly

Bairong Technology — Shenzhen, China
Data Scientist Intern, Financial Technology Department — *May 2020 – Aug. 2020*
- Built a semi-supervised learning model using the MixMatch algorithm with PyTorch to classify customers into different groups based on credit default risk predictions, achieving 90% accuracy
- Developed an automated report generation program to calculate performance indicators, graph statistical distributions, produce analytical summaries, and create formatted slideshows for clients

PROJECTS

Unsupervised Speech Recognition — *Apr. 2022*
- Implemented unsupervised audio-to-text transformation with the GAN-based wave2vec_U algorithm, enabling speech recognition for low-resource languages without sufficient training labels

Object Tracking System with Facial Recognition — *Mar. 2022*
- Constructed a target tracking system for both template matching and motion detection with the Lucas-Kanade method
- Incorporated CNN-based face recognition that achieved 82% accuracy on classification and 0.96 AUC on verification, providing an efficient and precise way to track down criminals with surveillance camera footage

Augmented Reality with Planar Homographies — *Feb. 2022*
- Conducted real-time image and video AR projections through interest point matching and homography estimation

Grocery Master — *Oct. 2021*
- Developed a software that allows users to search for a product, shows available options at nearby grocery stores, and compares their nutritional information by live-scraping Target, Walmart, and Trader Joe's websites

TECHNICAL SKILLS

Programming Languages: Python, Java, SQL, R, Stata, HTML
Tools: Cloud (Google Colab, Azure, AWS), ML Framework (PyTorch, scikit-learn, sktime, Pandas), Database (Hive, MySQL)