

# Hongling Lei

Apt 614, 6315 Forbes Ave, PA 15217

[honglinglei.github.io](https://honglinglei.github.io)

+1 (412) 805-8510

[www.github.com/HonglingLei](https://www.github.com/HonglingLei)

[hongling@andrew.cmu.edu](mailto:hongling@andrew.cmu.edu)

## EDUCATION

### Carnegie Mellon University

*Master of Information Systems Management, Business Intelligence and Data Analytics.* GPA 3.93 Pittsburgh, PA  
Aug. 2021 – Ongoing

- Relevant Courses: Deep Learning, Computer Vision, Data Structures for Application Programmers, Object Oriented Programming in Java, Unstructured Data Analytics (NLP), Machine Learning for Problem Solving

### Xiamen University

*Bachelor of Economics, Finance.* GPA 3.81 Xiamen  
Sept. 2016 – Jun. 2020

- Relevant Courses: Mathematical Analysis, Statistics, Probability Theory, Linear Algebra

### University of California, Berkeley

*Semester Exchange, Statistics, Computer Science, Business.* GPA 3.95 Berkeley, CA  
Jan. 2019 – May 2019

### Nanyang Technological University

*Summer Exchange, Digital Marketing.* GPA 5.0 Singapore  
Jun. 2018 – Jul. 2018

## PROFESSIONAL EXPERIENCE

### PPG Industries

*AI/ML Intern, Artificial Intelligence Center* Pittsburgh, PA  
May 2022 – Ongoing

- Build advanced machine learning models for customer relations, manufacturing, and supply chain processes
- Perform rapid experiments using AutoML and help deploy optimized models for production scaling

### Tencent

*Data Scientist Intern, Public Data Science Department* Shanghai  
Apr. 2021 – Jul. 2021

- Conducted causal inference with algorithms like Causal Bayesian Networks, X-Learner, and Causal Forest to analyze treatment effects of key drivers for customer consumption time on Tencent's streaming platforms
- Implemented experiments on user experience with Tencent WeSee, TikTok, and Kuaishou
- Constructed a data pipeline for automated causal inference, including modules like feature engineering, machine learning model training, and future experiment testing based on analytical results

### DiDi Chuxing

*Data Analyst Intern, Decision Support Department* Beijing  
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 alleviating traffic burdens in 14 Chinese metropolitan areas
- Created and monitored business dashboards that can automatically update data and visualize analytical results

### Bairong Technology

*Data Scientist Intern, Financial Technology Department* Shenzhen  
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

## PROJECTS

### Unsupervised Speech Recognition (GANs)

*Apr. 2022*

- Implemented unsupervised speech-to-text transformation with the wave2vec\_U algorithm, enabling speech recognition for low-resource languages without sufficient training labels

### Lucas-Kanade Tracking

*Mar. 2022*

- Developed real-time automated object tracking applications for both template matching and motion detection

### 3D Object Reconstruction Based on 2D Images

*Mar. 2022*

- Reconstructed 3D objects from stereoscopic image pairs using triangulation and the eight-point algorithm

### Facial Recognition (CNNs)

*Feb. 2022*

- Built a face recognition system that achieved 82% accuracy on classification and 0.96 AUC on verification

### Augmented Reality with Planar Homographies

*Feb. 2022*

- Conducted real-time image and video AR projections through interest point matching and homography estimation

## TECHNICAL SKILLS

**Programming/Software:** Python (PyTorch, Google Colab), Java, SQL, R, Stata, HTML, Apache Hive

**Languages:** Mandarin Chinese (native), English (fluent, TOEFL 116 with Speaking 30), Korean (beginner)