

# Jay Chiehen Liao

+886-988818967 | [jay.chiehen@gmail.com](mailto:jay.chiehen@gmail.com) | [linkedin.com/in/jc-liao](https://www.linkedin.com/in/jc-liao) | [github.com/jayenliao](https://github.com/jayenliao) | [rpubs.com/d920056](https://rpubs.com/d920056)

4+ years of experience in **Data Science & Machine Learning** • Proficient in **Python & R** programming • 10+ **AI/ML competitions**  
6+ years of experience in **Data Analysis & Statistics** • **Git, Linux, GPU, GCP** • 10+ academic presentations in various domains •  
Skilled in communication, model explanation, and problem-solving • **Interdisciplinary** backgrounds

## WORK EXPERIENCE

### **FAZZ FINANCIAL GROUP**

#### **DATA SCIENTIST (FULL-TIME)**

Sep 2022 - present | Taipei, Taiwan (Hybrid)  
- Focus on reconciliation and fraud & anomaly detection.  
- Help increase anomaly detection performance by 25%.

### **NETWORKS & AI LAB, NCKU**

#### **RESEARCH AND TEACHING ASSISTANT (PART-TIME)**

Aug 2020 - present | Tainan, Taiwan (Hybrid)  
- Research topics: Machine learning with graphs, graph neural network (GNN), deep learning for tabular data.  
- Managed **Linux** servers with **GPUs**.

### **HEALTH BEHAVIOR LAB, NCKU**

#### **RESEARCH ASSISTANT (PART-TIME)**

Jan 2017 - Jul 2020 | Tainan, Taiwan  
- Data analysis with **R** & **Python** for 5+ projects.

## EDUCATION

### **RENNES SCHOOL OF BUSINESS (RSB)**

#### **MASTER IN DATA & BUSINESS ANALYTICS (GPA=4.25/5)**

Sep 2021 - Apr 2022 | Rennes, France  
- Graduation Project: *Identifying Key Features of Standard Essential Patents in Information and Communication Technologies with Explainable Machine Learning*

### **NATIONAL CHENG KUNG UNIVERSITY (NCKU)**

#### **MASTER IN DATA SCIENCE (GPA=4.25/4.30)**

Aug 2020 - Aug 2022 | Tainan, Taiwan  
- The 1st Place, Academic Performance of The 1st Year  
- Thesis: *Learning Graph Structures from Tabular Data for Downstream Classification Tasks*

#### **DOUBLE BACHELOR IN STATISTICS AND PSYCHOLOGY**

Sep 2015 - June 2020 | Tainan, Taiwan

## SKILLS

### **LANGUAGES**

**ENGLISH** Professional working proficiency (**TOEIC 925/990**)

**MANDARIN/CHINESE** Native **FRENCH** Elementary

### **TECHNICAL SKILLS**

**PROGRAMMING** Python • R • Ruby on Rails **MARKUP LANGUAGES** Markdown • R Markdown •  $\text{\LaTeX}$  **VERSION CONTROL** Git  
**DEVELOPMENT ENVIRONMENT** VS Code • Jupyter • RStudio **OPERATION SYSTEMS** Linux • macOS Intel/M1 • Windows  
**MACHINE LEARNING** PyTorch • Scikit-Learn **GRAPH ANALYSIS** NetworkX • PyTorch-Geometric **CLOUD COMPUTING** GCP  
**DATA MANIPULATION** NumPy • Pandas • dplyr • tidyr • SPSS • MS Excel  
**DATA VISUALIZATION** ggplot2 • Matplotlib • Seaborn **DATA MANAGEMENT** MySQL • metabase2

## KEY PROJECTS & AWARDS

### **Predictive Modeling for 14-day Unplanned Hospital Readmission Risk by Using Machine Learning Algorithms BMC MEDICAL INFORMATICS AND DECISION MAKING (2021)**

- Constructed ML models with promising performance to predict patients' readmission.  
- Figured out influential features with model explanations.  
**MY ROLES:** Co-first author of the paper • Model constructor  
**SKILLS USED:** **Python** • **Catboost** • **Feature Engineering**

### **A Mobile App (UPrEPU) to Monitor Adherence to Pre-Exposure Prophylaxis in Men Who Have Sex with Men: A Protocol for a User-Centered Approach Mobile App Design and Development**

#### **THE JMIR RESEARCH PROTOCOLS (2020)**

**MY JOBS:** Conveyed public health rules with pseudo-codes to communicate with development engineers • Data cleaning  
**SKILLS USED:** **Pseudo-code** • **Python** • **R**

#### **TOP 1** out of 401 teams

### **CHUNGHWA POST BIG DATA COMPETITION 2019**

- Established ML models and proposed commercial solutions from the big data (>30GB) to reduce failed delivery rate (reduced about 65% eventually).
- Evaluated the proposed system in terms of ML performances, flexibility, and estimation of cost reduction (USD 376K and 44.8K hrs per year in Taipei City).

**MY ROLES:** **Leader** • **Model constructor** • **Presenter**

**SKILLS USED:** **Python** • **R** • **GBDT** • **Visual machines**

#### **TOP 3** out of 859 teams

### **RECOMMENDATIONS OF CREDIT CARD SPENDING**

#### **CATEGORIES 2021 (E.SUN BANK)**

**SKILLS USED:** **Python** • **Transformer** • **GBDT**

#### **TOP 25%**

### **AI CUP FALL 2020 - MEDICAL INFORMATION**

#### **DE-IDENTIFICATION (DEPT. OF COMPUTER SCIENCE AND INFORMATION ENGINEERING, NCKU)**

**SKILLS USED:** **Python** • **NLP** • **Transformer**

### **EXCELLENT AWARD** CERTIFICATE OF STATISTIC ANALYSIS (THE CHINESE APPLIED STATISTICS ASSOCIATION)

### **ADVANCED AWARD (TOP 3%)** COMPETITION OF STATISTICS (NCKU)