

Yu-Sheng (Kevin) Li

✉ kevin1kevin1k@gmail.com | 🌐 <https://github.com/kevin1kevin1k/> | 🔗 <https://www.linkedin.com/in/yu-sheng-li/>

INDUSTRY EXPERIENCES

Dcard, Taiwan

- Machine learning engineer, ads Mar 2022 – present
 - *Main goal: Maximizing ads revenue (with advertiser constraints) by optimizing online ads recommender systems*
 - Implement automatic weekly model re-training and deployment
 - Simplify ranking logic by removing 3 hard rules
→ CTR +12%
 - Survey and align offline evaluation metrics with online metrics
→ from mAP to gAUC and normalized entropy (scaled log loss)
 - Mentor and help onboard another MLE
 - Design a new A/B testing architecture for cost-per-click (CPC) bidding
→ from member-based to impression-based testing
 - Design and implement the CTR calibration framework for CPC ads
→ CPC goal achieve rate from 0% to 40% and then 80%
 - Improve LLM prompt engineering and design evaluation metrics for automatic ad title revision
→ adoption rate of revised version +50%
 - Feature engineering on contextual information (time, forum, position) and ad industries
 - Created a trouble-shooting dashboard for non-engineers
- Machine learning engineer, social May 2021 – Mar 2022
 - *Main goal: Maximizing user engagement by personalizing homepage post feed*
 - Rule-based hate speech traffic control
→ 10x reduction of daily hate speech comments
 - Remove hard rules and implement embedding-based candidate for post topic recommender system
→ overall topic selection CTR +8% and overall topics per post +14%
 - Improve the training set sampling strategy
→ homepage duration +0.7% and 10x distinct members in dataset (bias reduced)
 - Implement item-CF (EASE^R) recall candidate and replace three rule-based candidate
→ homepage PV +1.6%

Google, Taiwan

Jul 2019 – Sep 2019

- Software engineering intern
 - Identification and interactive clustering of test failure patterns for audio components

Alibaba Group, China

Jul 2017 – Sep 2017

- Software engineering intern
 - Training time reduction by 20%-30% of distributed training on parameter servers (which motivates Publication [3])

ACADEMIC EXPERIENCES

RIKEN & Tohoku University NLP Lab, Japan

Sep 2019 – Jan 2020

- Exchange research program (advisor: Prof. Kentaro Inui)
 - Context-aware revision for academic writing assistance

Teaching Assistant, National Taiwan University

- Formal languages and automata theory 2017 Fall, 2018 Fall, 2020 Fall
- Data structures and algorithms 2017 Spring, 2020 Spring

KDD Cup 2017, NTU Team Feb 2017 – Jun 2017

- Two tasks with final rankings **19/368** and **24/346**
- Contributions: Feature engineering, data analysis, and sequence to sequence learning

Academia Sinica, Taiwan Jul 2016 – Sep 2016

- Research intern (advisor: Dr. Wei-Yun Ma)
 - Question answering game with knowledge bases and online resources

**VOLUNTEER
EXPERIENCES**

Python Conference (PyCon) Taiwan

- Sponsorship team co-lead Oct 2022 – Sep 2023
- Sponsorship team member Mar 2022 – present

Taiwan Linguistics Olympiad Camp Jul 2019

- Lecturer on computational thinking

Traditional Chinese translation of the Python Documentation Aug 2023 – present

- Translator of parts of the Python 3.12 documents

PUBLICATIONS

- [1] Bowen Yuan, **Yu-Sheng Li**, Pengrui Quan, Chih-Jen Lin. Efficient Optimization Methods for Extreme Similarity Learning with Nonlinear Embeddings. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2021.
- [2] **Yu-Sheng Li**, Wei-Lin Chiang, Ching-pei Lee. Manifold Identification for Ultimately Communication-Efficient Distributed Optimization. In *International Conference on Machine Learning (ICML)*, 2020.
- [3] Wei-Lin Chiang, **Yu-Sheng Li**, Ching-pei Lee, Chih-Jen Lin. Limited-memory Common-directions Method for Distributed L1-regularized Linear Classification. In *SIAM International Conference on Data Mining (SDM)*, 2018.
- [4] **Yu-Sheng Li**, Chien-Hui Tseng, Chian-Yun Huang, Wei-Yun Ma. Guess What: A Question Answering Game via On-demand Knowledge Validation. In *International Joint Conference on Natural Language Processing (IJCNLP)*, 2017.

EDUCATION

National Taiwan University, Taiwan

- M.S. in Computer Science (advisor: Prof. Chih-Jen Lin) Sep 2018 – Jan 2021
 - Research and master thesis: Publication [1]
- B.S. in Computer Science, with minor in Mathematics Sep 2014 – Jun 2018
 - Undergraduate GPA: 4.08 / 4.3

SKILLS SETS

Programming Languages
 Python, SQL, C/C++, JavaScript, Shell scripting, Haskell

Skills and tools

- Data science and machine learning: scikit-learn, xgboost, PyTorch, TensorFlow, numpy, pandas, Jupyter Notebook
- Workflow orchestration and ETL pipelines: Airflow
- Database and feature stores: BigQuery, PostgreSQL, MongoDB
- Statistics: hypothesis testing and online A/B testing

COMMUNICATION

- Mandarin (native)
- English (fluent)
- Japanese (conversational)

HOBBIES

- Exercises (basketball, jogging, squash, bouldering, gym)
- Chess
- Puzzles