Shuangqi Li

Lausanne, Switzerland shuangqi.li@epfl.ch lishuangqi.com

Education

EPFL (Swiss Federal Institute of Technology of Lausanne) — *Lausanne, Switzerland*
PhD in Machine Learning — *Sept 2022 – present*

EPFL (Swiss Federal Institute of Technology of Lausanne) — *Lausanne, Switzerland*
Master in Data Science — *Sept 2020 – July 2022*

University of California, San Diego — *Remote*
Master in Computer Science (Quit due to visa issues and COVID-19.) — *Sept 2019 – June 2020*

University of Electronic Science and Technology of China — *Chengdu, China*
Bachelor in Microelectronic Science and Engineering — *Sept 2015 – June 2019*

Work Experience

Research Intern — *Zurich, Switzerland*
Oracle Labs — *July 2021 – Sept 2021*

- Developed a machine learning model that detects anomalous Linux sessions in the cloud servers.

Algorithm Engineering Intern — *Beijing, China*
Didi Chuxing (China's largest taxi-hailing platform) — *Oct 2018 – Feb 2019*

- Developed an algorithm for learning road segment weights from historical ride data, improving route planning quality for ride-hailing services.

Projects

Efficient Training Data Attribution for Large Language Models — *Aug 2025 – present*

Ongoing research

- Training Data Attribution faces severe computational challenges for modern-sized LLMs.
- We improved conventional influence functions by exploiting the low-rank property of the gradient to make them more efficient and scalable.

Improving Waste Detection and Sorting on Conveyor Belt — *Sept 2025 – present*
Supervising two semester projects: — *WasteFlow, Switzerland*

- YOLO uncertainty estimation and confidence recalibration.
- Real-time object brand recognition on conveyor belts.

Learning to Weight Parameters for Training Data Attribution — *Feb 2025 – Sept 2025*

Submitted to ICLR 2026. arXiv

- Identified the heterogeneity of attribution strengths across parameters/layers.
- Proposed a method to re-weight layers, boosting attribution accuracy and enabling fine-grained attribution.

Enhancing Text-to-Image Generation with Reliable Random Seeds — *Jan 2024 – Oct 2024*

ICLR 2025 Spotlight. arXiv

- Identified the significant role of initial noise in text-to-image inconsistencies.
- Proposed a method to mine reliable random seeds to improve text-to-image generation.

Controlling the Fidelity and Diversity of Deep Generative Models — *Feb 2023 – Mar 2024*

TMLR 2024 (poster presentation at ICLR 2025). arXiv

- Proposed an approach to bias deep generative models, such as GANs and diffusion models, towards generating data with either enhanced fidelity or increased diversity.

Interlock-Free Multi-Aspect Rationalization for Text Classification *Sept 2021 – Feb 2022*

Semester project. arXiv

- Proposed a multi-stage training method to alleviate the interlocking issue in selective rationalization.

Teaching Experience

- **CS-233: Introduction to Machine Learning** – *Head Teaching Assistant* *Spring 2025*
- **CS-401: Applied Data Analysis** – *Teaching Assistant* *Fall 2024*
- **CS-233: Introduction to Machine Learning** – *Teaching Assistant* *Spring 2024*
- **COM-407: TCP/IP Networking** – *Teaching Assistant* *Fall 2023*
- **COM-112: Object-Oriented Programming (in C++)** – *Teaching Assistant* *Spring 2023*
- **CS-456: Deep Reinforcement Learning** – *Student Teaching Assistant* *Spring 2022*

Honors & Awards

- **National Scholarship** *Sept 2018*
- **China Collegiate Programming Contest** – Gold Medal *May 2018*
- **China Collegiate Computing Contest** – First Prize *Mar 2018*
- **First-class People's Scholarship** *Dec 2017*
- **ACM ICPC (Asia Regional)** – Bronze Medal *Oct 2017*
- **China Collegiate Computing Contest** – First Prize *Apr 2017*
- **First-class People's Scholarship** *Dec 2016*

Skills

Programming: Proficient with Python, C++

Frameworks & Tools: PyTorch, Docker, Git, Linux, Cursor, Claude Code

Languages: Chinese (native), English (fluent), French (basic)

Publications

Learning to Weight Parameters for Data Attribution *arXiv*

Shuangqi Li, Hieu Le, Jingyi Xu, and Mathieu Salzmann *Submitted to ICLR 2026*

June 2025

Enhancing Compositional Text-to-Image Generation with Reliable Random Seeds *ICLR 2025*

Shuangqi Li, Hieu Le, Jingyi Xu, and Mathieu Salzmann *Spotlight (top 4%)*

Oct 2024

Controlling the Fidelity and Diversity of Deep Generative Models via Pseudo Density *TMLR 2024*

Shuangqi Li, Chen Liu, Tong Zhang, Hieu Le, Sabine Süsstrunk, and Mathieu Salzmann *Presented at ICLR 2025*

July 2024

Interlock-Free Multi-Aspect Rationalization for Text Classification *arXiv*

Shuangqi Li, Diego Antognini, and Boi Faltings *May 2022*