

LUPu Minh Long (Last Middle First)

minhlong9413@gmail.com |  [LinkedIn](#) |  [GitHub](#) |  [StackOverflow](#) |  [AI Exchange](#) |  [Google Scholar](#)

An innovative Data Engineer with 3.5+ Years of Experience in AWS, Snowflake Data Cloud, and SQL. Built robust Data Lake - Data Warehouse solutions using serverless services for Data Analytics, working directly with the VP, Director, Principal, and Senior engineers. A Computer Vision researcher with Professors and Ph.D. students from CMU, WISC, UIUC, and MBZUAI. An open-source contributor to a few of the most popular code bases. A top 5% yearly Stack Overflow contributor.

PROFESSIONAL WORK EXPERIENCE

FPT Software HCM

Data Engineer, now Data Architect

06/2021 - Present

Ho Chi Minh City

- Building a near real-time Data Lake - Data Warehouse solution from 10+ operational data sources with different formats to the Snowflake Data Cloud using AWS Serverless Services. The implementation reduced current costs by 91% and removed the dependency of a \$500k/year tool.
- Implemented AWS Lambda and Step Functions to extract and store data in S3 Buckets, monitored and tracked by CloudWatch and DynamoDB, with automatic Error Handling and Self-Healing in case of incidents and disasters.
- Analyzed and designed the data pipeline and data model for the Data Warehouse in the Snowflake Data Cloud, with Change Data Capture (CDC) and automatic ELT process, significantly cut down on human effort.
- Automated the Infrastructure setup and deployment process using Terraform with GitHub Actions CI/CD for reproducibility and faster deployment.
- Represented FPT Software to present the project with VP, Director, and Engineering leaders.
- Sponsored \$8000 by FPT USA Corp. to visit the customer's Headquarter in Atlanta, Georgia, USA.

ACADEMIC RESEARCH EXPERIENCE

CMU, WISC, UIUC, & MBZUAI

Machine Learning researcher

03/2021 - 04/2024

Remote

- Colleagues: [Haohan Wang](#) (AP @ UIUC, previously Ph.D. @ CMU), [Zeyi Huang](#) (Ph.D. student @ WISC)
Advisors: [Eric P. Xing](#) (Prof. @ CMU, President @ MBZUAI), [Yong Jae Lee](#) (Assoc. Prof. @ WISC).
Research focus: Image Classification, Object Detection & Segmentation, Out-of-Distribution Generalization, Transformers, Vision-Language models, and Multi-Modal LLMs.
- **Expeditious Saliency-based Mixup through Random Gradient Thresholding.** [Minh-Long Luu](#), Zeyi Huang, Eric P. Xing, Yong Jae Lee, and Haohan Wang. In [Practical-DL Workshop @ AAAI'23](#).
- **An Empirical Investigation Toward Understanding and Breaking the Bottleneck in Standard Experimental Protocol for Domain Generalization.** Zeyi Huang, [Minh-Long Luu](#), Yong Jae Lee, and Haohan Wang. 2022.
- **Benchmarking Object Detection Robustness to Unseen Domain Variations.** Zeyi Huang, Haohan Wang, [Minh-Long Luu](#), Eric P. Xing, and Yong Jae Lee. 2022.
- **RL-Mixup: Learning mix-up policies with Reinforcement Learning.** [Minh-Long Luu](#), Zeyi Huang, Haohan Wang. In [ML4Data Workshop @ ICML'21](#).
- **Reviewer** for International Conference on Learning Representations (ICLR) 2023, International Conference on Computer Vision (ICCV) 2023, European Conference on Computer Vision (ECCV) 2022.

AWARDS, ACCOLADES, & CERTIFICATIONS

- (2022) Snowflake SnowPro Core Certified ([418895447](#)), Udacity Nanodegrees: Data Scientist ([KGXSEDJL](#)), Data Engineering ([FXDYPHDN](#)), Data Analyst ([2NKMVM9E](#))
- (2020) Coursera Deep Learning ([YPXNKW8DLEHX](#)), Natural Language Processing ([7B73RNPYUB68](#))
- IELTS Academic (2022), Overall **7.5** (Read: 8.5, Listen: 7.5, Speak: 6.5, Write: 6.5) (equivalent to **TOEIC 950 / CEFR C1**)

PROGRAMMING LANGUAGES AND TOOLS

- Proficient Languages: Python, Git/GitHub, LaTeX, regex, SQL
- Familiar Languages: C/C++, JavaScript, HTML/CSS, AWS Cloud, Terraform, Java
- Business Intelligence tools: PowerBI, MicroStrategy
- Databases: PostgreSQL/MySQL/SQLite/MSSQL/Snowflake/Redshift, DynamoDB.
- AI/ML Libraries/Frameworks: TensorFlow, PyTorch, CatBoost, scikit-learn, RAPIDS
- Open-source contributions: [PyTorch](#), [torchvision](#), [DomainBed](#), [lightning](#), [accelerate](#), [TensorFlow](#), [cryptii](#), [NSA](#), [d2l-en](#)

EDUCATION

- Bachelor in Computer Science, International University - VNU HCM (09/2018 - 04/2022), 3.5/4 years graduation.