

Keifer Lee

ML RESEARCHER · COMPUTER ENGINEERING MAJOR

☎ (+1) 516-523-7477 | ✉ kl3866@nyu.edu | 🏠 <https://datacrisis.github.io> | 📄 www.github.com/datacrisis | 🔗 www.linkedin.com/in/keiferlee

Skills

PROGRAMMING	Python · Matlab · SQL · Bash · Assembly · Verilog · REST · Git · Solidity · Kubernetes · Docker
MACHINE LEARNING	PyTorch · Tensorflow · Scikit-Learn · Computer Vision · LLMs · 3D Vision · Robotic Perception · Generative AI
ANALYTICS	A/B Testing · DVC · DBT · Quarto · Snowflake · Redshift · OpenShift · Airflow · Elasticsearch

Experience

IBM New York City, NY
DATA SCIENCE INTERN May. 2023 - Aug 2023

- Developed a **MLOps-compliant NLP** starter-kit covering topics such Topic Classification via **IBM's Large Language Foundation Models**.
- Improved model accuracy by **+20%** and reduced development time from **2 weeks to 1 day**.
- Deployed** model via Red Hat's containerization platform, **OpenShift & Kubeflow**, demonstrating the ability to reduce deployment time by **90%**.
- Created a **Retrieval Augment Generation Q&A chatbot** with IBM's Foundation Model and other open-sourced model such as **LLAMA-2**.

New York University - AI4CE Research Lab New York City, NY
GRADUATE STUDENT RESEARCHER Oct. 2022 - Present

- Leading research project on optimal **novel view synthesis** via **Neural Radiance Fields (NeRF)** for autonomous robots and VR applications.
- Co-led an industrial project on **Sim2Real transfer** for robotic control; improved key metrics by **~96%** via **contrastive self-supervised learning**.
- Worked on very large-scale datasets and performed distributed training with **PyTorch** and **Slurm** on NYU's **supercomputer cluster**.

MoneyLion Inc Kuala Lumpur, MY
DATA SCIENTIST II Sep. 2020 - Aug 2022

- Led end-to-end projects** maximizing user acquisition funnel's return per unit risk taken with ML models and data-driven product feature tweaks, improving acquisition by **~100%** and **lowering credit risk** while improving system run-time by **~700%**.
- Led feature development projects including a pricing optimization tool, improving top-line metrics by **~17%**, and an optimal lead bidding algorithm (CAC vs LTV) for a high-priority lead generating channel with C-level stakeholders.
- Conducted daily experimentation work from exploratory analysis to large-scale statistical tests such as **A/B tests** for decision making.
- Created advanced **data visualization dashboards** via **SQL** for **Redshift** and **Snowflake** daily to provide near **real-time metric monitoring**.
- Led a large-scale (TB) database migration effort with Data Engineering using **DBT** and Codefresh **CI/CD** for seamless transitions.
- Worked with **MLOps** to establish **continuous model monitoring** and **automated retraining pipelines** for ML models. Consulted the team on procedures and key metrics to measure from both technical and product perspectives.
- Demonstrated effective knowledge transfer by **on-boarding 3 new team-members** onto complex projects as a mentor.
- Made significant contributions in time-sensitive incidence situations as part of the first-response task force, ensuring product stability, leading to multi-million dollars cost saved effectively.

Taylor's Unmanned Aerial Vehicle (UAV) Research Group Taylor's University, MY
ML RESEARCH ENGINEER Aug. 2019 - Apr. 2020
RESEARCH ASSISTANT Dec. 2018 - Aug. 2019

- Developed an **autonomous UAV** for **real-time monocular target tracking** with a light on-board single-board computer. **No pilots required**.
- Led the end-to-end development of a novel deep-vision algorithm for autonomous UAVs, from conceptualization, to deriving solutions, implementation and analysis. Result outperforms contemporary state-of-the-art model's accuracy with equal or higher computation efficiency.
- First-authored papers on ML with publications in the **American Institute of Physics Conference Proceedings**.
- Designed and trained various ML models with terabytes of data on both local machines and cloud servers via Bash.

Education

New York University New York, NY
MS IN COMPUTER ENGINEERING | CGPA: 3.78 Aug. 2022 - May. 2024

- Accepted and enrolled in the programme with a **\$14,000 scholarship**.
- Served as a **Graduate Research Assistant** @ AI4CE Lab on an industrial robotics control project for Spring 2023.
- Teaching Assistant** for the graduate **Robotics Perception** course for Fall 2023.

Taylor's University Subang Jaya, MY
B.ENG. (HONS) IN ELECTRICAL AND ELECTRONIC ENGINEERING | CGPA: 3.65 Mar. 2015 - Aug. 2019

- 5 times Dean's List and 2 times Book Prize Awardee.
- Awarded with the **Taylor's Grand Challenges Scholarship** (\$10,000) for academic excellence.
- Undergraduate research assistant with the Taylor's Unmanned Aerial Vehicle Research Group.

Projects & Extracurricular Activities

Leave Your Clothes Behind (LYCB)

New York University

COURSE PROJECT

Feb. 2023 - May. 2023

- Created a framework to convert video of clothing into virtual-wearable **3D garment models** for applications in AR and VR.
- Pipeline covered tasks such as **semantic segmentation**, **structure from motion (SfM)**, **novel view synthesis**, and **3D reconstruction**.
- Received full mark for the project in the course. Project can be found live at <https://github.com/IamShubhamGupto/LYCB>.

Cryptocurrency Arbitrage Bot

PERSONAL PROJECT

Aug. 2018 - Jun. 2019

- Wrote a **Python algorithm** to arbitrage Cryptocurrencies; e.g. Ethereum, Bitcoin, Solana. Capable of performing multi-step arbitrages between a chain of trading pairs.
- Developed a paper-trading (simulation) application with live feed data from various exchanges for real-time validation.
- Performed timing analysis and optimizations to reduce compute time by **~10x** with multiprocessing and asynchronous RESTful API calls.
- Could parse **400+** trading pairs across exchanges within **2 sec** and is extendable to most large-cap exchanges.

1stDayHack 2020

Selangor, MY

FOUNDER & LEAD ORGANISER

Sep. 2019 - Jan. 2021

- Organised and led **Malaysia's first Machine Learning Workshop+Hackathon** for **60 secondary students** from across Peninsular Malaysia.
- Created a beginner-friendly ML course, ran ML workshops, and oversaw the entire program.
- **Wrote and produced a free multi-chapter video based crash course**, covering topics from introductory Python to using advanced Deep Learning vision and NLP models.
- Created 1stDayKit, a ML toolkit that consolidating **12 state-of-the-art models** into one easy to use package. Written with **Python** and **PyTorch**.

Honors & Awards

2019 **Best High Impact Research Award**, 12th International Engineering Research Conference

Selangor, MY

2019 **Winner & Best Social Impact Award**, CodeathonX KL 2019

Malaya Uni, MY

2018 **Best High Impact Research Award**, 11th International Engineering Research Conference

Selangor, MY

2018 **Global Finalist & 1st Runner-Up**, NASA SpaceApp Challenge 2018

Kuala Lumpur, MY

References

Phai Teeratat Phongthiengtham, Ph.D.

phai@ibm.com

DATA SCIENCE LEAD & NLP DOMAIN FOCAL, IBM

Professor Chen Feng, Ph.D.

cfeng@nyu.edu

ASSISTANT PROFESSOR & HEAD OF AI4CE RESEARCH LAB, NYU

Stefa Etchegaray Garcia, Ph.D.

beetcheg@us.ibm.com

HEAD OF DATA SCIENCE, IBM

Andy Bryant

andy.bryant@ibm.com

SENIOR MANAGING STRATEGY CONSULTANT, IBM