

ZHIYU (EDWARD) LIANG

☎ 475-434-8259 ✉ edwardlzy0@gmail.com

EDUCATION

Yale University

Master of Science in Computer Science; **GPA: 4.0/4.0**

Aug. 2021 – May 2022

New Haven, CT

University of Toronto

Honours Bachelor of Science in Computer Science; **GPA: 3.84/4.00**

Sep. 2015 – Dec. 2019

Toronto, ON

TECHNICAL SKILLS

Languages: Python, Hack/PHP, JavaScript, Go, Swift, C/C++, Java, Shell, SQL, MATLAB, Markdown, Objective-C, Kotlin

Technologies: MERN, Hive, Presto, REST APIs, GraphQL, Kubernetes, Redis, Kafka, ELK stack, Bootstrap, Git, RegEx

Machine Learning: PyTorch, TensorFlow, Keras, NumPy, OpenCV, Fairseq, Tensor2Tensor, Scikit-Learn, Pandas

EXPERIENCE

Meta

Software Engineer

Jul. 2022 –

Menlo Park, CA

- Working on the **Product Detail Page** team of Facebook and Instagram Shops.
- Drove 10+ teams collaboration to make **Shops on Facebook website** compliant with EU regulations (DMA & DSA). Technologies used include React, GraphQL, JavaScript and PHP.
- Jointly developed "preview mode" to PDP rendering that resulted in up to **80% TTRC (Time To Recent Content) improvement** on high variant count products, and **dropped PDP abandonment rate by 5.5%**.
- Implemented efficient checkout page pre-fetching that accounted for **20% faster TTI (Time To Interactive)**.
- Led various **better engineering** and **operational excellence** efforts such as improving the test coverage by 40% and driving weekly SEV reviews.
- Maintained 30+ performance and business metrics logging events and **reduced 5%-10% logger payload size**.

Qualcomm

Machine Learning Engineer

May 2018 – Apr. 2019, Feb. 2020 – Jul. 2021

Toronto, ON

- Developed 2 quantization algorithms to reduce the size of **Transformer models** by **4x** while preserving **97.7%** model accuracy in **machine translations**.
- Proposed and trained **GAN-based image-denoising algorithm** that achieved state-of-the-art performance without requiring pixel-aligned training data.
- Implemented a Python script to translate internal model representations into ONNX format using Regular Expression, which accelerated the processing time **from 2 hours to 5 seconds**.
- Developed a Python **web crawler** that collected more than 550,000 code snippets for training a Large Language Model for programming languages.
- Jointly developed the Quantization-Friendly MobileNet, won the **1st prize in 2018 IEEE Low Power Image Recognition Challenge** and published a **NeurIPS Workshop paper**.

Vector Institute

NLP Research Intern

May 2019 – Dec. 2019

Toronto, ON

- Crawled and preprocessed large-scale high-quality text data of **12 billion word tokens**.
- Implemented the **GPT-2 model** in Tensor2Tensor and training scripts for **distributed training** setup, available for hundreds of researchers and sponsors to use.
- Profiled distributed training performance of GPT-2 on **256 GPUs**, which helped decide the multimillion-dollar purchase of additional training hardware.

PROJECTS

Basic Shops PDP | Hack/PHP, JavaScript, React, Objective-C, Kotlin

- Built the Basic Shops Product Detail Page on Facebook and Instagram app and website for EU and UK users that complies with EU regulations (DMA & DSA).
- Safeguarded and monitored modules and data presented to hundreds of millions of visits from the EU every day.

FaceBlock App | Tensorflow, Java, Android Studio | 2018 Qualcomm Hack Mobile Winner

- Developed an Android app using Java and TensorFlow to protect people's privacy in live video streams by detecting, tracking and replacing unwanted faces with a selected emoji in real time on Samsung Galaxy S9.
- Won 1st place in 2018 Qualcomm Hack Mobile Hackathon** out of 60+ teams and 250+ participants.

Travel Planner Web App | TypeScript, MERN Stack, Google Maps API, Jest, Netlify

- Developed and deployed a travel planning website that automatically generates a travel itinerary from a selected city and dates. The tourist attractions and restaurants are recommended based on Google Maps ratings.