Brian Li

► b77li@uwaterloo.ca · J 647-523-8026 · 🏶 https://brianljx.vercel.app/ · 🛅 Brian Li

Skills

Programming Languages: Python, C++, Java, HTML/CSS, Tailwind CSS, Javascript, Typescript, SQL, Dart

Frameworks: React, NextJS, NodeJS, Flutter, Langchain, Pytorch, Tensorflow, Keras, Selenium, Jenkins Tools and Technologies: Git, Linux, AWS, Docker, Kubernetes, STM32, Arduino, Raspberry Pi, OpenCV, FPGA,

FFmpeg, Whisper, ElasticSearch, Figma

RELEVANT EXPERIENCE

Eon Media

AI/ML Developer Intern

April 2024 - August 2024

- Spearheaded research and integration of NLP capabilities into **OpenSearch cluster**, performed testing and deployment of **TorchScript models** and **ingest pipelines**. Utilized models for natural language understanding (NLU) of input text and conversion of metadata into **dense vectors** for **KNN-indexing**.
- Designed and trained a custom **binary classifier** from scratch to identify logos in images using **Convolutional Neural Networks** (**CNNs**) and **transfer learning** of the **ResNet50 model**. Achieved **93**% accuracy, improving existing classifier's performance by **45**%.
- Improved text detection model through introduction of dewarping function, leveraging a **Generalized Additive Model (GAM)** to calculate curvature and map points to perpendicular offset using Bresenham's algorithm; resulted in a **50**% improvement in recall rate and an **80**% reduction in runtime of pod.
- Facilitated **cloud deployment of pods** through **EKS clusters**, performed and optimized **containerization** of various pods and algorithms, reducing **docker image** sizes by **40%** and **resolving bottlenecks in pipeline.**

Wat.AI

Lead Neural Network Developer

September 2023 - Present

- Led sub team of 3 in designing and prototyping **sparse** and **denoising autoencoders** using **PyTorch** for compression of IoT cybersecurity data (CICIOT 2023); optimized draft autoencoders with **exponential learning rate decays** and **learning curve analysis**
- Contributed in writing and updating <u>team substack articles</u> to highlight progress of team project; helped with the testing of **machine learning models** focused on cyberattack detection for IoT devices

Waterloo Formula Electric

Lead Firmware Developer

September 2023 - Present

- Created hardware-in-the-loop (HIL) tests using **Python** to validate electric car components and determine expected behavior of unit and **RTOS**; utilized tools such as **STM32IDE**, **Virtual Box** and **Vagrant** to find translation between code and firmware input values
- Analyzed the Battery Management Unit (**BMU**) by stimulating its state of charge; conducted in-depth examination of **multithreading** firmware code in **C** and schematics to identify source variables and functions

PROJECTS

"Spotify-Roots" — NextJS, Typescript, Flask, Langchain, llama.cpp, Qdrant \mathbf{Q}

- information of
- Developed a web application using a **RAG-based approach** to provide origin information of song, album or artist, integrating Genius API for annotation extraction and **Llama 3.1 model** for inference and conversion of data.
- Application built with OAuth2 protocol for Spotify account logins, allowing users to perform analysis on personal playlists.

- Created a software that visualizes **real-time** data sheets in various graphic forms; data sheets are pulled from OurWorldinData with seamless **processing** of complex data
- Visualizations are **dynamic** and customizable based on user inputs and adjustments

"Sumo Car Bot" — Python, Arduino, Raspberry Pi, HTML, Linux, Flask

- Led a group of 4 in developing a fully operational miniature car capable of **autonomous navigation** and user control via a **Flask** hosted web server; organized and planned individual tasks for members
- Utilized **Arduino** and **Raspberry Pi** for the implementation of the embedded system; coded car functions within a **Ubuntu** environment using **SSH protocol** and **VIM**. Constructed with WHIMIS guidelines

EDUCATION

University of Waterloo