Joshua Zhang

jzhang0224@gmail.com | joshuazhang.vercel.app | linkedin.com/in/joshzh | github.com/Teddygat0r

Education

University of Washington Expected Graduation: June 2027
GPA: 3.93, Bachelor of Science in Computer Science *Seattle, WA*
Relevant Coursework: Distributed Systems, Data Structures & Algorithms, Systems Programming

Experience

Stripe - *Software Engineering Intern* June 2025 - Current

- Created a daily AWS S3 export of tax information fetched from a source of truth, enriching it with tax rates and time dimensional information to surface inconsistencies and improve data accuracy and SOX compliance.
- Led MongoDB migrations for Stripe's tax eligibility data from broad product-level mappings to granular, atomic-level classifications, reducing inconsistencies by 99% and ensuring alignment with up-to-date tax rates.
- Developed an agentic bot using RAG for self-service tax analysis, allowing non-technical teams to query tax data directly from our data stores without seeking engineering help.

Amazon - *Software Engineering Intern* Jan 2025 - June 2025

- Enabled seamless integration between legacy ASP.NET (C#) systems and modern Amazon Books frameworks.
- Used table pagination and server side database processing to achieve a 98% increase in speed in a Brilliance Publishing internal tool.

Art of Problem Solving - *Software Engineering Intern* June 2024 - Sept 2024

- Developed an online placement exam using React used by over 20,000 students across 12 campuses.
- Built a Node.js middleware REST API to relay frontend requests to a legacy PHP backend.
- Enabled gradual deprecation of the PHP server by decoupling it from the frontend and PostgreSQL database.

Advanced Robotics at UW - *Controls SWE* Sept 2023 - Sept 2024

- Maintained Taproot, an open-source embedded systems framework written in C++ for Robomaster robots.
- Built a real-time Python-OpenCV system for red square detection, 3D pose estimation, and serial transmission of robot orientation data.

Department of Homeland Security - *Software Engineering Intern* June 2023 - Aug 2023

- Developed a Python-based RAG system with a local LLM and vector database, allowing users to query lengthy documents for targeted insights.
- Integrated ChromaDB with a local LLM via LangChain to enable retrieval-augmented generation, improving response accuracy and reducing hallucinations by 30%.
- Modeled the LLM API after the OpenAI API, allowing it to communicate with the LangChain tool.

Projects

Spotify Recommender | *Python, Flask, Scikit-learn, NextJS, React, Typescript, Tailwind*

- Generated vector embeddings for 12,000 songs to identify their similarity and create recommendations in Python.
- Built a frontend using React, giving users a graphical visualization of songs and their recommendations.

WildEye AI | *Flask, Tensorflow, NextJS, React, Typescript, Tailwind, Python*

- Built a wildlife-spotting camera powered by an NVIDIA Jetson Nano that automatically snaps photos and identifies animals that wander by.
- Trained a model using Tensorflow to identifying 90 different animal species with an 86% test accuracy.

Technical Skills

Languages: C++, C#, Typescript, Javascript, NodeJS, Python, Java, SQL, HTML/CSS, Bash
Frameworks: React, PyTorch, Spring, Express, ASP.NET
Developer Tools: Git, Microsoft Azure, AWS, Docker, Kubernetes
Interests: Distributed Systems, Distributed Model Training, LLM Inference, Mixture of Experts, Database Systems