William Yang

will.yang.j25@gmail.com · GitHub · **LinkedIn** · **Personal Site** · 339-927-8810 · Boston, MA

EDUCATION

Massachusetts Institute of Technology — **Cambridge, MA**
Master of Engineering in Computer Science | GPA: 5.0/5.0 — *Jun. 2026*

Massachusetts Institute of Technology — **Cambridge, MA**
Bachelor of Science in Computer Science and Mathematics, Minor in Writing | GPA: 4.9/5.0 — *Jun. 2025*

TECHNICAL SKILLS

CS Coursework: Grad Distributed Systems, Grad Database Systems, Grad Deep Learning, Operating Systems, Compilers, Advanced Algorithms, Advanced Software Design, Grad Computer Vision, Grad Natural Language Processing, Web Development
Languages: Python, C++, Go, Java, Typescript, JavaScript, C, Rust, Ruby
Frameworks: ReactJS, Node.js, JUnit, Mockito, Jest, PyTorch, NumPy, MatplotLib, SciPy

RESEARCH

MIT Rus Lab — **Cambridge, MA**
Machine Learning Researcher — *Sep. 2024 – Jun. 2025*

- Worked on an algorithmic and drone swarm system to tag whales in live ocean environments, researched published in **ISER 2025**.
- Trained and optimized a vision model to identify bounding boxes of whales in live drone footage images.
- Built an end-to-end drone simulation in Python where a drone swarm used vision algorithms to target and then land on a pod of whales.
- Created a new point cloud registration algorithm to accurately assign drones to an individual whale to land on in ocean settings

MIT Fiete Lab — **Cambridge, USA**
Machine Learning Researcher — *Jan 2024 – Jun 2024*

- Worked on building linear, self-attention layers, and a positional embedding for a discrete token transformer model in PyTorch
- Built Cuda Kernel software in Numba and Triton for both forward and backward passes of the discrete transformer
- Final paper published in the Neural Information Processing Systems journal

MIT Camera Culture Lab — **Cambridge, USA**
Machine Learning Researcher — *Jan 2023 – Dec 2023*

- Built a **Neural Radiance Field** model to use satellite and drone images to generate 3D views of cities, detect shadows, and track moving objects within scenes. With a novel model architecture, I improved accuracy by **15%** in 3D rendering.
- Used **Blender** and Google Maps **API** to generate **5000** satellite images with varying sun angles and camera heights. Created an integrated pipeline standardizing images from varying satellite image formats and lighting conditions.
- **Collaborated** with undergrad and grad researchers to **write final paper** accepted to the 2024 **EarthVision Conference**

PROFESSIONAL EXPERIENCE

Stripe — **Seattle, WA**
Software Engineer Intern — *Jun. 2025 – Present*

- Worked on developer productivity team to speed up user requests **by 3 times** for data engineering **AWS machines** for 700 engineers.
- Wrote cron jobs in **Puppet/Ruby** that maintained standby pools of **300 remote hosts** for GPU and high compute data machines.
- Enabled instance storage caching on data engineering remote machines, speeding up git operations and build times by **5 times**
- Created update workflows to maintain and update packages/repos for a pool of **250 EBS** and attach them to remote hosts on user request.

Amazon — **Seattle, WA**
Software Dev Engineer — *June 2024 – Aug 2024*

- Transferred data for **7 million** fashion product items to a **SQL** searchable **AWS Open Search** database from **DynamoDB**
- Reduced search latency times on fashion products internal database by **ten-fold** through optimizing AWS tooling
- Built out **3 APIs,** test suites, and developed image storage and transfer infrastructure on **AWS CDK, S3, and Lambda**
- Integrated **full stack** features in **Node.js** and **ReactJS** for a customer feedback and reviews for fashion products on the stores site

NLmatics — **Cambridge, MA**
Machine Learning Engineer Intern — *Jun 2022 – Feb 2023*

- Led the intern team in writing a paper summarizing our model's results, linked here [**link**]. Developed an ensemble model version of **BERT** on **AWS** to perform question answer tasks on various documents from legal, medical, and academic settings, beating baseline models by **10 percent in accuracy**.

PROJECTS

Go Database – [link]

- Built a fully function SQL database in Go consisting of a query parser, file storage, and a buffer pool cache storage system
- Implemented a BTree data structure to store database elements, **tripling** typical database query speeds
- Integrated database scheme onto various cloud computing frameworks such as **AWS Lambda, Spark,** and **Snowflake**

C++ Poker Bot – [link]

- Built a C++ poker bot to play a variant of Texas Hold'Em for the MIT PokerBots competition using counterfactual regret minimization. My team's bot ended up **placing second** among C++ poker bots in the final tournament.