Steve Li

steveli@college.harvard.edu
github.com/lithafnium
http://steve-li.com
978-735-3386

Education:

Harvard University '23 Cambridge MA
Relevant Coursework: Systems Programming and Machine Organization, Data Structures and Algorithms**,** Machine Learning, Probability, Linear Algebra and Differential Equations, Theory of Computation

Technical Skills:

Java, C/C++ , Javascript (React.js, React-Native, Vue.js, Node.js, Electron.js, Jest, Typescript), Python, Bash, Flask, MySQL, PostgreSQL, Firebase, GraphQL, GCP, CI/CD, AWS (Amplify, DynamoDB, S3, Lambda, EC2, ECS, EBS), Docker, HTML/CSS

Relevant Experience:

Gamalon | Machine Learning Engineer Intern Cambridge MA, Sept 2021 - Present

- Currently developing a novel **neural network architecture** based on probabilistic factor graphs using **PyTorch**, implementing Log-Sum-Exp activation functions as a ReLU proxy using trained log mass probabilities.
- Designing a suite of **CLI** tools using **Bash**, **Python**, and **GPT-3** for model building and question/answer generation, reducing manual business analyst labor from **~4 weeks to ~4 days**
- Creating question/answer generation prototypes using **nltk n-gram clusterin**g and **topic hierarchie**s with google-searched background information to supplement client chat-bot knowledge.

Amazon | Software Development Engineer Intern Seattle WA, June 2021 - Aug 2021

- Developed public **REST API** for resource tagging for Amazon Connect, AWS's cloud call center, designed to be shipped to **hundreds of thousands** of customers by September.
- Spearheaded redesign of resource cleanup **lambda** workflows and architecture, reducing user throttling rates by over **110%**
- Implemented access control support using tagging through IAM roles, allowing for resource allocation and user restriction.
- Debugged **production** code and wrote **unit/integration tests** required for **CI/CD pipelines** to maintain AWS cloud resources.

Fractal | Software Engineering Intern Cambridge MA, Jan 2021 - May 2021

- Created testing infrastructure with **~10x** fewer bugs to ensure quality software using **Jest, React-Testing-Library, and Enzyme,** complete with **Github Actions workflows**.
- Implemented AWS resource tracking with **logz.io docker** integrations and slack notifications through **Github Actions workflows** to monitor **EC2 instances** and **ECS clusters,** saving **thousands of dollars** in sunk costs.
- Developed **Electron.js app** and **Flask Python webserver** functionality with **e2e** and documentation.

Pine Park Health (YC S18) | Software Engineering Intern Berkeley CA, May 2020 - Sept 2020

- Created **frontend** UI elements to display user tasks and patient symptoms, increasing **productivity** and ensuring patient safety using Vue.js for Toggbook, an application aimed at managing medical staff in assisted living communities.
- Implemented **REST API** using PostgreSQL and Node.js, to manage and store **thousands** of patients and staff over **21** health-care facilities. Improved automation and monitoring by creating a streamlined data platform connecting a central store to the database, saving **days** of manual labor.

Software Projects:

Wikipedia Topic Modeling (Top2Vec, Typescript, Flask)
Designed a topic visualizer for Wikipedia articles using **Top2Vec** and **transformer-based NLP models** trained on multilingual data. Created **flask** backend and **typescript custom graph editor** too**l.**

Spotify Curator (Python, Flask, Pandas, Scikit-learn, React.js)
Devised a song preference predictor by analyzing playlists and liked songs on Spotify. Utilized a random forest classifier to examine empirical song data from a user to predict whether they would like an inputted song or not. Tuned to perform at 75% accuracy.

Personal Website (React.js, Typescript)
Personal website built completely from scratch, compiled using Snowpack and written using Typescript/React
http://steve-li.com

Shell (C++)
Created a full-fledged command shell with functionality for **fork**, **exec**, **interrupts**, **pipes**, **background processes,** and **redirects.** Pipelines were implemented using **fork**, **dup2**, and **close** in order to manage file descriptors, as well as interruption with process group control.

Clubs and Activities:

Wave Learning Festival | Technical Co-Director California, April 2020 - Present
Built frontend **React.js** app with **e2e** testing used by over **10000** students **worldwide** in more than **60** countries. Redesigned technical workflow by creating **CI/CD** infrastructure on **AWS Amplify** servers with **DynamoDB** Backend, **S3** storage, **Lambda** function deployments, and **GraphQL.**

Harvard Undergraduate Data Analytics Group | Analyst, Associate Cambridge MA, Sept 2020 - May 2021
Co-lead a team of six members on a Data Science research project for an Athletic Clothing brand. Developed metrics based on user-data and customer insights to provide valuable action items.