Sachin Iyer

sachin@sachiniyer.com

https://sachiniyer.com

https://github.com/sachiniyer

Education

New York University Sep 2020 — May 2023

Bachelors of Science in Computer Science *Brooklyn, NY*

Experience

AWS Bedrock Software Development Engineer *Seattle, WA 12/2023 — Present*

Bedrock: Creating Generative AI infrastructure for AWS Bedrock. Part of the core team and launched/maintained key features like batch inference, model distillation, and provisioned throughput with AWS-Wide and ReInvent visibility.

NYU High Speed Research Network (HSRN) Academic Researcher *Brooklyn, NY 02/2021 — 05/2023*

Parallel File System: Deployed an NSF funded 6PB storage PFS (SeaweedFS) for usage internally and externally to the HSRN. Automated Deployment with Ansible Playbooks and Rust CLI. Benchmarked with Bonie++ and IOR.

Audio Conferencing: Created an audio service (Portaudio) in C++ that interfaces with internal broker service.

CI/CD: Developed documentation, linting, testing, deployment (DinD with Kaniko) pipelines for the project in Gitlab

Mentorship: Leader of the student research arm. Managed and onboarded over 110 students over 4 semesters

Amazon Last Mile Software Development Engineering Intern *Seattle, WA 05/2022 — 08/2022*

Data Aggregation Service: Created a full stack service to visualize last mile delivery data. Created a Typescript React frontend with Polaris styling that calls an AWS backend implemented with Java Lambda Functions, API Gateway, S3, and Internal Amazon Services. Used S3 Select and Spark to filter through about 1TB per query.

Hewlett Packard Enterprise (Aruba Networks) Cloud Intern *Santa Clara, CA 06/2019 — 08/2019*

Estimating Bandwidth: Estimated bandwidth using Auto ARIMA/Prophet and other time series algorithms.

Sparkup Software Development Engineering Intern *Brooklyn, NY 09/2022 — 12/2022*

UX Development: Implemented a new feature in React Native App to link names and phone numbers in transactions

Dark Forest Graphics Intern *San Jose, CA 12/2021 — 02/2022*

Shader Development: Created a typescript plugin that allows for custom WebGL shaders in the Dark Forest game.

Projects

pong-wasm: A RL model for pong trained with policy gradients built with candle, compiled to wasm, and runs in your browser. Utilized web features like IndexedDB, Web Workers and custom serialization with hacks into wasm-bindgen.

K3S Cluster: Created a portable resilient fault-tolerant k3s cluster that networks through a wireguard mesh (headscale). Wrote openresty lua to filter by SNI and stream Proxy Passed TCP. Created custom monitoring, uptime and a wiki

mnist-wasm: A framework-less (just ndarray), resource efficient, rust based wasm model that trains/predicts in your browser. Built with just in time spawned web-workers, the rust-based yew frontend framework, axum, wasm-bindgen/gloo

Invoice Categorization: Automatically sorts invoices into categories with Bedrock Batch, Flask, React, Tailwind. Wrote a preprocessing step with a common crawl adapter to supplement data about companies. Also dogfooded my service.

Wikipedia Editor: Used PySpark, Cohere Embeddings, NYU HPC/ SLURM to analyze 6TB of dumps. Found metrics on the breath of topics wikipedia editors touch with LDA and the variance of topics they edit.

Apps Status: Built an API with Rust, Tokio Async, Axum, and Reqwest that proxies status for my self-hosted apps through a Rust Lambda function with a custom Megalodon-rs that pushes outage statuses to Mastodon every 5 min

Sembox: A drive with semantic searching over documents with blip, xsum, whisper, bert, mui/nextjs, by taking the cosine similarity of summaries/search terms. Supports images, text, video, audio through intelligent document type recognition

Tweet Toxicity: Used DistilBERT, Pytorch, HuggingFace Transformers, Streamlit and AdamW to classify toxicity type

Resow: Created a better Craigslist free section with ReactJS, MUI, Open Layers, Express, MongoDB, S3, Mocha, Jest

Ansible Batch Runner: Used Rust and Clap to create a cli for batch running and managing Ansible Playbooks

Skills

Courses: Deep Learning Systems, Machine Learning Systems, Operating Systems, Computer Networking, Linear Algebra

Langs: Typescript, Javascript, Node, C++, C, Rust, Java, Python, (e)Lisp, Bash/Zsh, Cuda, Kotlin, Perl, Lua, LaTeX

Certifications

AWS Solutions Architect *September 2021*

AWS Cloud Practitioner *August 2021*

Stanford Machine Learning by Andrew Ng *July 2020*