

SAMUEL YUAN

samuelyuan.com • syuan@utexas.edu • (346) 857-9528 • linkedin.com/in/YuanSamuel • github.com/YuanSamuel

EDUCATION

The University of Texas at Austin B.S. Computer Science Honors, BBA Business Honors, BBA Management May 2025
Overall GPA: 3.89/4.0

EXPERIENCE

Databricks – Software Engineering Intern; Mountain View, CA May 2024 – August 2024

- Spearheaded TCP/HTTPS networking connectivity probe service to monitor health of 550+ critical cloud resources globally
- Established multitenant thread pool/worker model to efficiently schedule and execute 35K+ probes hourly across all shards
- Implemented dynamic self-onboarding schema and CLI with customizable Prometheus alerting and routing per probe

Roblox – Software Engineering Intern; San Mateo, CA May 2023 – August 2023

- Orchestrated system to store and serve 500M+ searches from 150M users in CockroachDB through gRPC endpoint in .NET
- Pioneered chi-squared algorithm to detect trends from 40M+ daily searches, enhancing user insights from data warehouse
- Architected and implemented Spark, Hive, and Airflow-based system to generate and store trending search candidates daily
- Led migration from workstation to server garbage collection to decrease memory usage by ~5% and stabilize CPU usage

Striveworks – Software Engineering Intern; Austin, TX May 2022 – August 2022

- Constructed Golang microservice to analyze and report product dependencies, deployed on Kubernetes cluster with Helm
- Designed Kubernetes-in-Docker cluster to deploy local instance of microservice for testing, mocked out various features

RESEARCH

UT Machine Learning for Systems Lab – Undergraduate Researcher January 2024 – Present

- Lead effort to implement joint learning in the OS to predict more accurate cache and prefetching decisions in real time
- Formulate attention-based contrastive learning encoder model to learn representations of system traces, e.g. cache features
- Evaluate PyTorch and C++ LibTorch model, integrating encoder into SoTA cache/prefetcher models to improve performance

UT NLP and Computational Linguistics Lab – Undergraduate Research Assistant September 2024 – Present

- Productionize CoditT5 code LLM by optimizing model performance and training efficiency and deploying on HuggingFace
- Fine-tune and benchmark model, achieving up to 19% performance improvement over SoTA for bug fixing and code reviews
- Direct group discussions and provide analysis on cutting-edge LLM research for code understanding and generation

National AI Institute for Foundations of Machine Learning (IFML) – Undergrad. Researcher November 2023 – February 2024

- Applied Momentum Contrast encoder model to ESM2 embeddings to detect residue-level protein sequence similarity
- Implemented Triplet model with InfoNCE loss to compare training with mined positive and negative pairs to MoCo model
- Integrated PyTorch models with preprocessed MSA dataset and evaluate similarity of aligned and non-aligned residues

PROJECTS

PicTag – Founding Engineer December 2021 – Present

- Built AI-driven semantic search engine for mobile and web platforms leveraging CLIP model for image relevance using Flutter
- Engineered Golang API to create image vector embeddings from S3 bucket using AWS Lambda and SQS, stored in vector DB
- Pitch and prepare product demo to dozens of businesses and NASA researchers, garnering 50+ app waitlists for launch

VS Code GitHub Projects – Full Stack Developer February 2021

- Released open-source VS Code extension providing real-time interface for GitHub Projects with 25,000+ total downloads
- Established persistent GraphQL communication by polling between multiple WebViews rendered by Svelte and GitHub API

LEADERSHIP EXPERIENCE AND ACTIVITIES

Honors Business Association – Vice President of External Affairs (Fall 2022) Fall 2021 – Spring 2023

- Led committee of 3 to organize various philanthropic events involving 30+ members to serve the local Austin community
- Organized food drive with honors program office collecting 25+ canned goods donated to the Central Texas Food Bank

HONORS

- Neo Scholar Finalist August 2023
- Top 3 at UT Machine Learning and Data Science Datahack February 2022
- First Place at UT Computer Science Honors CritterFest AI Competition October 2021

ADDITIONAL INFORMATION

Programming Languages: Go, Python, CUDA, Java, JavaScript, TypeScript, Dart, C, C#, Rust, Scala, SQL

Frameworks and Platforms: PyTorch, Spark, Hive, Airflow, GraphQL, Node.js, .NET, React, Swagger, Pandas, Kubernetes