Shelby Yang

Austin, TX | github.com/Wilye |linkedin.com/in/shelbyyang | shelbyyang.com | shelbyyang@utexas.edu

EDUCATION

The University of Texas at Austin May 2026
Turing Scholar
B.S. in Computer Science, Minor in Economics
Relevant Coursework: Honors Data Structures, Honors Algorithms, Honors Operating Systems, Honors Artificial Intelligence, Machine Learning, Honors Discrete Math, Honors Computer Architecture, Linear Algebra, Probability

SKILLS

Experienced in: Python, Java, C/C++, Swift, Supervised Learning, Unsupervised Learning, Reinforcement Learning, TensorFlow, XGBoost, Scikit-Learn, Pandas, NumPy, SQL, Spark, R, Golang, Apache Thrift, GraphQL, React, UNIX, Verilog
Familiar with: Apache Solr, Docker, Kubernetes, Apache Kafka, HTML/CSS, JavaScript

EXPERIENCE

Google - *Software Engineer Intern, Gmail Security* August 2025 - November 2025

- Engineered a DNS crawling pipeline using C++ and Flume on Borg to automate the collection of 10+ record types (MX, SPF, DKIM), replacing manual "dig" command workflows for abuse investigations.
- Implemented a recursive crawling algorithm to map complex dependencies across CNAME targets, SPF includes, and PTR lookups, effectively clustering networks of related abusive domains.
- Designed a time-series Spanner schema to store historical DNS data, creating a foundational dataset used to identify security misconfigurations and serve as a critical signal for Gmail's anti-spam models.

Apple - *Software Engineer Intern, CoreOS Power and Performance* May 2025 - August 2025

- Integrated a predictive tool for estimating available memory on Apple devices into analytics systems, enhancing visibility into memory usage trends.
- Implemented an end-to-end data pipeline to process 10,000 user powerlogs using Spark, enabling large-scale statistical analysis of memory consumption across workloads.
- Provided deeper insights into workload memory patterns to guide future product decisions.

Reddit - *Software Engineer Intern, Search Relevance* May 2024 - August 2024

- Developed "Query Results Comparison System", an internal tool that increased visibility for how a query gets processed throughout Reddit's backend services (Go, Python).
- Used React and Javascript to develop the GUI and make HTTP requests. Wrote GraphQL queries and modified schemas to hydrate the candidate IDs returned with additional fields like post titles and subreddit names.

PROJECTS

Animal Shelter Placement Prediction (Python)

- Designed and implemented an end-to-end supervised learning pipeline to predict animal shelter outcomes, including data cleaning, feature extraction, and model evaluation.
- Used feature engineering, SMOTE, PCA, hyperparameter search, and model tuning techniques.
- Achieved 0.7979 balanced accuracy with a Random Forest model.

Operating System Kernel Development (C++)

- Designed and implemented a custom operating system kernel and support for system calls, emulated on QEMU.
- Features include an EXT2 file system, interrupt and exception handling, cooperative and preemptive multithreading, garbage collection, offline threads handling, virtual memory, and synchronization primitives.

LEADERSHIP & COMMUNITY INVOLVEMENT

UTCS Directed Reading Program (DiRP) - *Administrator* August 2022 - Present

- Organized DiRP, an organization aimed at promoting undergraduate engagement in computer science research.
- Oversaw a membership of 500+, managing 17 distinct research groups spanning areas such as architecture, graphics, robotics, machine learning, systems, and theory.

Texas Undergraduate Computational Finance - *Analyst* January 2023 - Present

- Wrote biweekly market analysis reports and presented investment strategies using quantitative techniques. Strategies focused on predictive modeling for stock performance and risk management.