Justin Zheng

510-737-8367 | jzheng394@gatech.edu | linkedin.com/in/justinzheng9398 | github.com/justin5764

Education

Georgia Institute of Technology Expected May 2026

Bachelor of Science in Computer Science | GPA: 4.0/4.0 *Atlanta, GA*

Relevant Coursework: Data Structures & Algorithms, Linear Algebra, Computer Organization & Programming, Probability & Statistics with Applications, Systems & Networks, Artificial Intelligence, Database Systems

Experience

Software Development Engineer Intern May 2025 – Present

Amazon *Tempe, AZ*

- Developed a Seller Central homepage card to help brands enrolled in Amazon's Brand Referral Bonus program claim millions in unclaimed referral credits, leveraging AWS Lambda, DynamoDB, and Amazon SQS
- Increased visibility of $23MM in unclaimed credits for ∼131,000 merchants by deploying optimized daily ETL jobs
- Accelerated tax form status updates as measured by immediate UI changes, using event-driven AWS infrastructure
- Reduced internal API traffic by ∼50% by streamlining backend processes through asynchronous data handling

Undergraduate Researcher January 2025 – Present

Georgia Institute of Technology - CHART Lab *Atlanta, GA*

- Developed BrainBridge, a collaborative application using Python, Flask, and React, to enhance interdisciplinary team communication through real-time AI-powered speech and language processing
- Engineered and integrated advanced AI/NLP pipelines, including transformer models, vector embeddings, and text classification, to bridge domain-specific language gaps
- Implemented real-time speech-to-text transcription and speaker diarization using WhisperX and pyannote.audio, achieving 90%+ accuracy

Machine Learning Researcher June 2022 – August 2022

The Pennsylvania State University - Professor Suman Saha *Remote*

- Researched and implemented machine learning models for credit card fraud detection using algorithms such as Logistic Regression, Decision Tree, and Random Forest
- Analyzed and visualized dataset of 284,807 credit card transactions using pandas, Matplotlib, and seaborn to identify patterns in fraudulent behavior
- Optimized model performance by identifying key features from imbalanced datasets, resulting in improved accuracy and F1 scores, achieving up to 99.951% accuracy with XGBoost

Projects

Virtual Memory Simulator | *C*

- Developed a comprehensive virtual memory simulator in C, implementing core OS functionalities like virtual-to-physical address translation, page table management, and demand paging
- Improved performance on diverse workloads (A*, MCF, Perlbench) using FIFO, Approximate LRU, and Random page replacement, reducing page faults by up to 42.9% and AMAT by up to 16170 clock cycles
- Optimized disk I/O by up to 32.5% via efficient dirty page tracking and write-back mechanisms

Multithreaded CPU Scheduler | *C*

- Architected a C-based CPU scheduler (FCFS, RR, SRTF, Priority Aging), reducing average wait times by up to 22%
- Governed process lifecycles and thread safety using pthreads, mutexes, and condition variables for robust concurrency
- Furnished a CLI for dynamic algorithm/parameter selection, enabling comparative performance analysis across workloads

Workout of the Day (WOD) Prediction | *Python*

- Developed a Python ML pipeline predicting CrossFit Open scores using athlete data, workout text, and benchmarks
- Achieved 0.57 test MAE (4.6% improvement) by tuning Random Forest and PyTorch neural network hyperparameters
- Enhanced accuracy via LLM-based workout description embeddings (PCA) and robust feature engineering

Technical Skills

Languages: Java, Python, C, C++, HTML/CSS, JavaScript, TypeScript, SQL
Frameworks: Next.js, React, Node.js, Express, Tailwind CSS, FastAPI
Tools: Git, MongoDB, Firebase, AWS, MySQL, Jupyter