

Albert (Geyang) Xu

858-568-6771 | albertxu1010@gmail.com | [LinkedIn](#) | [Portfolio](#) | Los Angeles, CA

EDUCATION

University of California, Riverside <i>Ph.D. in Computer Science (Incoming)</i>	Sep 2026 - Expected <i>Riverside, CA</i>
University of California, San Diego <i>M.S. in Computer Science (GPA: 3.88/4.0)</i>	Sep 2022 - Mar 2024 <i>La Jolla, CA</i>
University of Liverpool <i>B.S. in Computer Science (GPA: 3.93/4.0, top 3%)</i>	Sep 2018 - Jun 2022 <i>Liverpool, UK</i>

PUBLICATION

PVLDB'25: SAVAGE — Stress-Testing ML Pipelines with Adversarial Data Corruption (2nd author). [[arXiv](#)] · [[Code](#)]

EXPERIENCE

Fuyao Glass America Inc. <i>Software Engineer</i> - Built a YOLOv8 computer-vision pipeline for inline automotive-glass inspection, detecting scratch, chip, edge, and coating defects on line-camera frames - Integrated detections into Python/SQL ETL with coating/tempering/inspection data; shipped role-based dashboards for OEE, FPY, scrap, and changeover - Productionized CV inference + ETL with cron, alerts, and access control — 99.3% job success and 2.4 hrs/week less downtime; supported A/B and DOE setpoint studies in Jupyter	Oct 2025 - Apr 2026 <i>Moraine, OH</i>
4Pexonic Inc. <i>Software Engineer</i> - Developed an academic impact analysis platform tracking citation patterns across tens of thousands of records using Node.js, MongoDB, and Next.js with Docker, serving university researchers and policy institutions - Built two-tier caching (in-process LRU + Redis), cutting median and p95 latency by 40% - Integrated gRPC microservices; implemented bcrypt/session auth and Nginx TLS	Oct 2024 - Oct 2025 <i>San Diego, CA</i>
UCSD Halicioğlu Data Science Institute <i>Research Engineer</i> - Built modular Python Injector pipeline (generate → sample → inject → eval) with beam-search + Optuna/TPE, cutting downstream AUC by >0.25 vs. random attacks - Co-authored SAVAGE (PVLDB'25), designing corruption dependency graphs and bi-level black-box search to model mechanism-aware missingness, selection-bias, and outlier patterns for pipeline-level stress-testing of ML systems - Ran Inject → Clean → Retrain benchmarks across missing-value, selection-bias, and outlier scenarios to evaluate state-of-the-art cleaning, debiasing, and UQ pipelines and expose their data-centric robustness gaps	Oct 2023 - Oct 2024 <i>La Jolla, CA</i>

PROJECTS

SurfStore: Distributed Cloud Storage Server Project Link - Designed and implemented a Dropbox-like cloud-based file storage service using Go, featuring distributed storage and consistent hashing algorithms for efficient file management - Built for large-scale data handling, the project showcased excellent scalability during tests and efficiently met the data storage demands of thousands of users through client-server communication via gRPC - Enhanced the system's fault tolerance by implementing the RAFT distributed consensus protocol, which led to a 50% improvement in error recovery efficiency during testing

TECHNICAL SKILLS

Programming Languages: Python, Java, Go, SQL, C/C++, JavaScript, R

Frameworks and Tools: Docker, Git, gRPC, Node.js, Next.js, Django, React, MongoDB, Optuna, Jupyter, YOLOv8, AWS