

Yuming Xie

+1 908 258 1623 | taox0001@outlook.com | [GitHub](#)

EDUCATION

Brown University

M.Sc. in Computer Science (GPA: 4.0/4.0)

Concentration: Software Engineering, High-Performance Network Systems, Formal Verification

Providence, RI

Aug 2024 – May 2026

Rensselaer Polytechnic Institute

B.Sc. in Computer Science & Mathematics (GPA: 3.72/4.0)

Concentration: Algorithms, Machine Learning, Mathematical Models of Optimization & Operations Research

Minor: Quantitative Modeling of Economics

Troy, NY

Aug 2020 – May 2024

SKILLS

Java, Python, Rust, C++; AWS (Lambda, ECS Fargate, SQS, CloudWatch, CDK), Docker, Linux, CI/CD; MySQL, PostgreSQL, Redis, OpenSearch; Spring Boot, Flask, .NET; SonarQube, Clang-Tidy, ASan/Valgrind; Data Structures & Algorithms, OOD, Complexity.

SELECTED PROJECTS

Multi-LLM Dual-Mode Debugging System (DMDS)

- Built a cloud-local debugging pipeline (Python/Flask, queued job-runner, Git-history context) orchestrating verify-feedback loops across Qwen/Llama to detect and ship fixes end-to-end.
- Delivered two modes (1) cloud high-throughput screening (Bandit, Clang-Tidy, SonarQube) and (2) local deep checks (static + dynamic + formal + complexity) via CLI for broader automated analysis; **p95 static screen <30 s; deep-check median ~3 min** (LLM/API-bound).
- On 100 IBM Project CodeNet tasks, reduced a composite failure metric (40% static + 30% dynamic + 20% formal + 10% complexity) by **27.8%** (Python), **22.4%** (Java), **14.3%** (C++) over feedback cycles.
- Migrated long-running batches from Lambda to ECS Fargate (CDK) to bypass the **15-minute** Lambda cap; added standard retries and error handling to improve batch success rate.
- Implemented case-based retrieval: after bug categorization, query a MySQL index of prior fixes (reuse-on-hit / persist-on-miss) to standardize patches and reduce median time-to-resolution on reuse hits.
- Co-authored a VS Code extension to trigger cloud static runs and local deep checks from the editor.

SchemaBridge — Prompt-Guided Excel Normalizer

- Built a Spring Boot service that ingests non-relational Excel + user prompts and outputs schema-validated relational JSON, including header normalization and type/constraint inference using JSON Schema.
- Streamed parsing of very large .xlsx via Apache POI XSSF Event/SAX (XSSFReader) to keep memory bounded; standardized/normalized cell types pre-validation to reduce parse-time errors and variance.
- Designed request/response-validated REST endpoints to ease integration and testing across clients.
- Implemented embedding-assisted header mapping: generated embeddings (DJL + Hugging Face) and indexed them with Amazon OpenSearch k-NN (HNSW/FAISS) for approximate k-NN lookup, improving header-match accuracy on **80+** files vs string-match baseline.

Decentralized Finance — Large Transactions Model (DeFi LTM)

- Built a CI/CD generate-verify pipeline for Solidity (Remix compile, property-based tests, Mythril scans, Etherscan ingestion) with audit trails; achieved **64%** functional, **42%** security pass-rates over **400+** test cases.

PUBLICATIONS

- Xie, Y.** (2024). *Comparison of Various Recommendation Algorithms and Statistical Analysis of Datasets*. In Proc. CONF-SPML 2024.
- Xie, Y.**; Nishizawa, T.; Zhang, H.; Chen, J.; Ker, A. (2024). *Self-Supervised Learning for Predicting Breast Cancer Treatment Response Using Public Datasets*. Accepted at BSSL 2024.

ACADEMIC EXPERIENCE

Rensselaer Polytechnic Institute

Teaching Assistant, Algorithms

Troy, NY

Aug 2022 – Dec 2022, Aug 2023 – May 2024

- Led weekly recitations/office hours (~30 students); authored solution outlines and set up autograding.