

Xinyi Huang

☎ (672)551-9217 | ✉ xinyi_huang@sfu.ca | [in](#) Xinyi Huang

Education

Simon Fraser University

Ph.D. in Computer Science | Supervisor: Zhengjie Miao

2024 – Present
Burnaby, BC, Canada

- GPA 4.00/4.33
- Research area: RL for data science, agentic RL
- Mitacs Globalink Graduate Fellowship

Sichuan University

B.Eng. in Computer Science and Technology

2019 – 2023
Chengdu, China

- GPA 3.89/4.00
- Graduated with Honors

Internship Experience

Microsoft

Research Intern | Mentor: Haoyu Dong

May 2025 – Aug 2025
Beijing, China

- Post-trained tabular QA model via GRPO in a multi-node, multi-GPU environment using **verl** and **Ray**.
- Developed a multi-stage agent pipeline with LangGraph to handle large spreadsheet QA tasks.
- Built specialized modules for dynamic context selection, data preview and query verification, allowing the agent to navigate and synthesize information from tables that were too large to encode fully.

Mitacs (DataPrep.ai)

Research Intern | Mentor: Jiannan Wang

Jun 2022 – Sep 2022
Burnaby, BC, Canada

- Developed a module to generate performance reports with Gantt charts, CDF charts, and tables to diagnose slow SQL queries.
- Visualized statistics like table scan history to help identify and remove unnecessary indexes.
- Implemented error detection for common query issues (e.g., missing columns) before compilation.
- Implemented query tuning techniques to optimize slow query bottlenecks.

Preprints

SPA: A SQL-Plan-Aware Reinforcement Learning Framework for Query Rewriting with LLMs Submitted to VLDB 2027

- Developed a plan-aware GRPO reward strategy, enabling Qwen-32B (LoRA) to outperform GPT-5.4 and Gemini-2.5-Pro in query rewriting by incentivizing plan change over superficial lexical changes.
- Extended GRPO with multi-stage rewards, with **probability-gated adaptive reward** for query-level curriculum learning.
- Designed adaptive reward shaping and **on-policy self-improvement** to mitigate sparse rewards, reduce slowdown rewrites, and improve runtime/tail-latency on IID and OOD workloads.

Structure-Grounded Knowledge Retrieval via Code Dependencies for Multi-Step Data Reasoning

- Engineered a novel retrieval framework that organizes domain knowledge via a graph induced by **function-call dependencies**, specifically for complex code generation tasks.
- Implemented a dependency-path retrieval logic that identifies critical links between semantic input/output tags, providing a more precise context extraction than embedding-based similarity.
- Validated performance on DABStep, demonstrating that grounding retrieval increases accuracy by 109.1%-283.3% over traditional and graph-based RAG methods while retrieving substantially less context.

EVE-Bench: Diagnostic Evaluation of LLMs' Capabilities in SQL Reasoning

- Developed EVE-Bench, an execution-grounded framework for SQL equivalence verification, utilizing counterexample generation as a diagnostic lens to pinpoint LLM reasoning bottlenecks.
- Utilized schema abstraction to identify LLMs' potential reliance on semantic hints, detecting when models bypass structural reasoning in favor of lexical cues.
- Benchmarked 8 LLMs across 2,900+ query pairs, discovering that the most advanced models outperform traditional formal methods by up to 155% while exhibiting complementary failure modes.

Projects

Develop Children Autism Diagnostic Scale with Bayesian Factor Analysis

2022

- Extract children's behavioral characteristics and key diagnostic features from patient cases and diagnosis.
- Designed an autism diagnostic scale using Bayesian Factor Analysis to assist diagnoses.