

Junliang Lyu

Dallas, TX | +1-972-890-5140 | junliangl@smu.edu | [linkedin.com/in/junlianglyu-mscs](https://www.linkedin.com/in/junlianglyu-mscs) | github.com/Junliang-Lyu

EDUCATION

Southern Methodist University (SMU)

Dallas, TX

Master of Science in Computer Science, Expected Dec 2026 | GPA: 3.40 / 4.00

Relevant Coursework: Machine Learning in Python, Algorithm Engineering, Database Systems, Software Engineering

TECHNICAL SKILLS

LLM Systems: Conversational AI, RAG, long-term memory, multi-stage LLM workflows, model routing, prompt engineering; OpenAI-compatible APIs, Claude API, Ollama

ML & Retrieval: GraphCodeBERT, CodeBERT, FAISS / vector search, Transformers, PyTorch

Programming Languages: Python, Java, JavaScript, SQL (MySQL), HTML/CSS

Frameworks & Tools: AstrBot, Streamlit, Django, Flask, React, Docker, Git, GitHub

Backend & Systems: REST APIs, MCP (Model Context Protocol), agentic tool/skill interfaces, Backend Development, Layered Architecture

Databases: MySQL (schema design, query optimization, data validation), SQLite

EXPERIENCE

IT & Network Support — Mingjin Furniture Co., Ltd.

Mar 2023 - Jun 2023

Dongguan, Guangdong, China - 500-person manufacturer (~60 computer-equipped office staff)

- Wrote Python scripts to automate the finance team's recurring spreadsheet workflows — consolidating multi-department Excel reports and distributing them — removing repetitive manual merging.
- Co-managed company-wide IT across ~60 workstations: operating-system deployment, network setup and configuration, and day-to-day hardware/software troubleshooting alongside one other technician.

PROJECT EXPERIENCE

Multi-Channel AI Chatbot Platform (Independent Project)

2026

- Built a multi-channel conversational AI platform with layered moderation, persona RAG, long-term memory, and model routing; production-tested on QQ in a 5,000+ member community (5,689 requests over 22 days) and validated end-to-end on Discord.
- Designed a safety and interaction funnel with allow/deny controls, harassment interception, lightweight/local intent classification, and progressive enforcement before long-context generation, reducing unnecessary long-prompt calls.
- Built a project-specific web data collection and knowledge-preparation pipeline for persona-grounded retrieval; integrated persistent user memory plus time-aware schedules and randomized character events.
- Resolved SQLite write-lock contention under 300+ active-user load using WAL mode and targeted PRAGMA tuning; containerized services with Docker Compose, health checks, local-only admin interfaces, and internal network isolation.

AI Code Search Engine (Research Project with SMU Faculty)

2025

- Built a semantic code search engine using GraphCodeBERT embeddings and FAISS approximate nearest-neighbor indexing over a 300K-snippet subset of CodeSearchNet, enabling natural-language retrieval by semantic similarity.
- Designed a modular preprocessing -> embedding -> FAISS indexing -> retrieval pipeline with an MRR/Recall evaluation harness, benchmarking 7 retrieval approaches under one consistent setup.
- Reached MRR 0.50 on a held-out test set, versus 0.31 for CodeBERT and 0.09 for a token-overlap baseline evaluated identically.

House Price Prediction - Data Mining (Team Project)

2026

- Built an end-to-end ML pipeline on the Ames Housing dataset (2,930 samples), using correlation analysis, variance filtering, and mutual-information selection to reduce 138 raw features to 107 predictors.
- Benchmarked Lasso, Random Forest, and XGBoost via 5-fold cross-validation with grid search; XGBoost reached $R^2 = 0.93$ (RMSE = \$22,947), reducing prediction error 32% versus the Lasso baseline.
- Identified Overall Quality and Total Square Footage as dominant predictors (76% combined XGBoost feature importance) and applied a log transform to stabilize variance across price ranges.

ADDITIONAL INFORMATION

Awards: Dean's List; University Merit Scholarship

Work Authorization: F-1 STEM OPT (eligible for up to 3 years of work authorization; no H-1B lottery required during the OPT period)