Ryan Zhang

New York, NY | rz2716@columbia.edu | (917) 929-8106 | rhyzhang.github.io | rhyzhang | rhyzhang

Education

Columbia University — *New York, NY*

Bachelor of Arts in Computer Science | John Jay Scholar (Top 10%) | GPA 4.01 — *Sept 2024 – May 2028*

- **Coursework:** Multivariable Calculus, Probability & Statistics, Data Structures, Discrete Math, Advanced Programming.

Experience

Researcher & ML Engineer — *New York, NY*

Machine Learning for Good Lab, NYU Courant — *July 2023 – present*

- Researching algorithmic bias by applying a custom Bayesian model to learn and understand judges' decision making to predict sentencing outcomes.
- Designed and implemented the majority of the research ML pipeline (model implementation, data collection, training, visualizations) from a PhD student's calculations.
- Increased model speed by 60x by linearizing key functions, improving experimental iteration speed by 10x.
- Created complex synthetic datasets and proved model convergence, a critical step in the research process.
- Training model on the real-world dataset on HPC; working towards publishing a paper.

Data Scientist & Manager — *New York, NY*

Rebecca & Christopher Cleaners | Project Case Study: Dry Cleaning DS — *Jan 2024 – present*

- Architected an end-to-end automated ETL pipeline, migrating the business from intuition based operations to data driven decision making by consolidating fragmented data sources.
- Reverse engineered undocumented legacy POS schemas (SQL Server, MDB) and developed custom scripts to extract, sanitize, and load sales data into a centralized data lake.
- Developed a Streamlit analytics dashboard to visualize KPIs and revenue patterns, identifying opportunities during the holiday season to optimize operations scheduling.
- Developing a churn prediction model to identify at-risk high-value clients and performing market basket analysis to drive service bundling strategies.
- Responsible for customer service, managing employees, handling finances and solving real-world problems; worked 12 hour shifts, 6 days a week (summer 2024).

Co-Director — *New York, NY*

Columbia AI Alignment Club, Columbia University — *Jan 2025 – present*

- Secured $50,000 in competitive grants to fund technical research tracks, enabling 150+ students to access high level AI safety resources and mentorship.
- Facilitated technical analysis of frontier LLM capabilities, leading advanced reading groups on mechanistic interpretability, RLHF, and scalable oversight.
- Selected for exclusive retreats and symposiums (OASIS @ Constellation, UChicago), directly engaging with researchers from Anthropic and top labs to understand emerging risks in foundation models.

Projects

HTTP Web Server & Database Engine — *Nov 2025 – Dec 2025*

- Technologies: C, BSD Sockets, TCP/IP, Linux, Make, Valgrind
- Architected a concurrent HTTP 1.0 web server and a standalone database server from scratch in C, utilizing BSD sockets for robust client-server communication.
- Implemented a reverse-proxy architecture to handle static file serving and dynamic database queries, ensuring strict memory safety and zero leaks via Valgrind.

Wave Generator — *Jan 2023 – May 2023*

- Technologies: C, Arduino, Python, Streamlit, Robotics, Git
- Built the entire control stack from scratch in C for a Wave Generator
- Created a GUI interface that communicates to the Arduino wirelessly using MQTT protocol
- Wired the motors, motor controller, and Arduinos into one compact system

Technical Skills

Languages: Python, SQL, Java, C, HTML/CSS
Data Science: pandas, NumPy, scikit-learn, Matplotlib, Seaborn, Streamlit
Machine Learning: Regression, Classification, Algorithmic Fairness/Ethics
SWE: pytest, Valgrind, Make, TCP/IP, HTTP
Tools: Git, Docker, Conda, UV, Linux/WSL, JupyterNotebooks, HPC, Agentic Programming, GenAI/LLMs