

Stella (Juanru) Zhang

New York, NY • (607) 280-8686 • stllazhn@gmail.com
linkedin.com/in/stella-zhang-891321254 • github.com/stllazhn

EDUCATION

New York University, New York, NY

M.S. in Data Science, Sep 2026 – Expected May 2028

Cornell University, Ithaca, NY

B.S. in Information Science, Aug 2022 – May 2026

Relevant Coursework: Machine Learning, Natural Language Processing, Data Mining, Database Systems, Information Retrieval, Algorithms

TECHNICAL SKILLS

Programming Python, SQL, R, Java, JavaScript, C++, OCaml, HTML, CSS

ML & Data scikit-learn, XGBoost, random forest, gradient boosting, logistic regression, k-means, PCA, cross-validation, hyperparameter tuning, feature engineering, MySQL, SQLite, LangChain, retrieval-augmented generation (RAG), text-to-SQL, prompt engineering, large language models (LLMs)

Analytics and Visualization pandas, NumPy, Power BI, Tableau, Matplotlib, seaborn, ggplot2, Microsoft Excel, A/B testing, statistical analysis

Developer Tools GitHub, Jupyter, VS Code, Node.js

Languages English (native), Mandarin Chinese (native), French (basic)

RESEARCH EXPERIENCE

Cornell University — Human-LLM Value Alignment Project (Ithaca, NY)

Research Assistant to Prof. Aditya Vashistha, May 2025 – Present

- Design an experimental evaluation framework on ConvoKit's CRAFT model to measure how LLM-generated constructive comments alter the rhetoric of online discussions.
- Analyze large-scale discussion corpora in R and SQL, producing evidence that model-generated interventions homogenize the cultural and ideological values expressed by human speakers.
- Co-author *LLMs Homogenize Values in Constructive Arguments on Value-Laden Topics*, submitted to ACM CHI 2026, documenting systematic value misalignment and cultural flattening in LLM-rewritten text.

PROFESSIONAL EXPERIENCE

Convooloo (Remote, US)

Software Engineer Intern, NLP, Jun 2024 – Oct 2024

- Built a natural-language-to-SQL system with LangChain RAG and Google Gemini at 95% accuracy on patient health data.
- Integrated MySQL and SQLite with AI inference pipelines for real-time query execution across heterogeneous sources.
- Implemented conversational retrieval chains with custom templates rendering SQL output as human-readable responses.

Xiaomi (Beijing, China)

Data Analyst Intern, Jun 2024 – Aug 2024

- Validated weekly product data across 5,000+ SKUs, auditing labels, costs, revenue, and inventory status for reporting accuracy.
- Built 4 Power BI dashboards on product lifecycle, regional sales, and inventory turnover to guide pricing and restocking.
- Streamlined SQL and Excel reporting workflows across 6 regions, cutting report construction time by 30%.

Hisense (Qingdao, China)

Software Engineer / Data Analyst Intern, Jun 2023 – Aug 2023

- Developed and implemented debugging protocols for a traffic statistics tracking and classification system.
- Integrated Alibaba Sentinel to rate-limit user data requests, mitigating service overload and cascading failures.

Baidu — Virtual Data Analysis Mentorship Program (Remote Work, China)

Data Analyst / Programmer, Mar 2023 – May 2023

- Benchmarked XGBoost, Random Forest, Logistic Regression, and Naïve Bayes on 10,000+ web visits, reaching 95%+ accuracy.