

Hannah Yimei Zhao

hannahzhao2026@u.northwestern.edu | 1016 Garnett Pl, Evanston, IL 60201
949-590-7571 | www.linkedin.com/in/hannah-zhao-

EDUCATION

Northwestern University

Evanston, IL

Dual Degree: Bachelor of Arts in Statistics and Data Science; Bachelor of Music in Violin Performance

September 2022 – June 2026

- Cumulative GPA: 3.945/4.00; Dean's List 8/8; SAT: 1580/1600
- Relevant Coursework (Completed): Data Structures & Algorithms in Python, Data Management and Information Processing in SQL, Advanced Machine Learning in Python, Data Science in R, Regression Analysis
- Activities: Kappa Phi Lambda Sorority Vice President & Community Service Chair; LingMech Lab Research Assistant; Northwestern University Contemporary Music Ensemble Violinist
- Awards: Northwestern Wildhacks Hackathon (2024: Crowd Favorite, 2025: Finance Track 1st Place)

PROFESSIONAL EXPERIENCES

Seashine Capital

Irvine, California

Finance & Accounting Intern

January 2025 - Present

- Assist in the month-end close process by reconciling accounts and ensuring the accuracy of financial records.
- Utilize Excel and Google Sheets to organize financial data, create reports, and contribute to financial analysis.

AutoNavi (Alibaba)

Beijing, China

Data Analyst Intern

July 2023 – August 2023

- Utilized SQL to query and analyze live map data for quality and accuracy from Li Auto's automated driving technology.
- Built an interactive Alibaba Deep Insight dashboard (like Tableau) to monitor mapping data efficiency and accuracy.

PROJECTS

Deep Learning for CMIL Tissue Patch Classification

- Fine-tuned a DenseNet121 CNN model in PyTorch, modifying classifier head and applying transfer learning techniques.
- Designed a full preprocessing pipeline with torchvision transforms, resizing patches to 224×224, applying ImageNet normalization, and addressing fine feature preservation challenges in small medical images.
- Implemented class imbalance strategies (Random Oversampling, Undersampling, and Class-Weighted CrossEntropyLoss) and built a Multiple Instance Learning (MIL) structure to group patch data by case.

Faith, Behavior, and Mental Health in University Students

[Link to Blog Post](#)

- Cleaned and organized large survey data (124 columns) in Python (Jupyter Notebook), reducing to 28 key columns grouped by demographics, faith, behaviors, and mental health.
- Utilized wireframing (Mokkup.ai) to plan dashboard structure, employing Dynamic Zone Visibility and filters.
- Published a blog post (Medium) summarizing findings and actionable insights to demonstrate data storytelling and analytical impact.

Predicting Goodreads Ratings Based on Book Characteristics

[Link to Github](#)

- Conducted thorough data preprocessing on a dataset of 52,748 books, incorporating feature engineering (e.g., imputation, encoding, and interaction terms) and transformations to improve model accuracy.
- Implemented and compared multiple machine learning algorithms (Random Forest, XGBoost, ensemble models, etc.).

TECHNICAL SKILLS

- **Programming Languages:** Python, R, SQL
- **Machine Learning & AI:** Supervised/Unsupervised Learning, Neural Networks, TensorFlow, scikit-learn, Gradient Boosting (XGBoost, LightGBM), Clustering, Natural Language Processing (NLP), Model Evaluation, Cross-Validation
- **Data Analysis & Preprocessing:** Feature Engineering, Data Cleaning, Dimensionality Reduction (PCA, UMAP)
- **Tools & Software:** Jupyter Notebook, VS Code, Git/GitHub, Tableau, Microsoft Excel
- **Languages:** English (fluent), Chinese (fluent)

ADDITIONAL INFORMATION

Musical Activities: Civic Orchestra of Chicago Violinist, Napa Valley's Summer Music Festival Assistant 2nd Violin Principle

Interests: Long-Distance Running, Content Creation (TikTok and YouTube), Christianity, Thrifting, Reading