

# Linh Tran

[ltran18@u.rochester.edu](mailto:ltran18@u.rochester.edu) | [linkedin.com/in/linh-ph-tran/](https://www.linkedin.com/in/linh-ph-tran/) | [github.com/lnhtrn](https://github.com/lnhtrn)

## EDUCATION

### UNIVERSITY OF ROCHESTER

Rochester, NY

*MS in Computer Science, Artificial Intelligence Specialization*

*Anticipated 12/2025*

### UNIVERSITY OF ROCHESTER

Rochester, NY

*BS in Computer Science, Minor in Mathematics*

*12/2024*

- **GPA:** 3.5/4.0, Dean's List, Schwartz Discover Grant for Undergraduate Research, Davis UWC Scholarship.
- **Coursework:** Machine Learning, Deep Learning, Data Structures & Algorithms, Data Mining.

## SKILLS

- **Programming Languages:** Python (advanced), Java (advanced), SQL, R, JavaScript, CSS, HTML
- **Toolkits:** Jupyter Notebook, Tableau, PowerBI, Snowflake, Microsoft Azure, Google Cloud Platform, Linux
- **ML/DS Libraries:** PyTorch, Pytorch Lightning, Transformers, NumPy, Pandas, Matplotlib, Scikit-learn, Wandb

## WORK EXPERIENCE

### Dye Lab, University of Rochester

Rochester, NY

*Data Science Research Assistant*

*09/2024 – Present*

- Utilized large language models in Python to generate keywords, cluster, and classify 10,000+ global COVID survey responses, enhancing healthcare insights on ethnic group socio-economic positions.
- Collaborated closely with the social science team, demonstrating strong communication and teamwork.

### kLab, University of Rochester

Rochester, NY

*Deep Learning Research Assistant*

*01/2024 – Present*

- Implemented experience replay techniques with Python and PyTorch Lightning to reduce racial bias in 20+ facial recognition models, improving accuracy across all racial groups by 10-15%.
- Trained and fine-tuned 20+ DCNN models, achieving a minimum biased accuracy of 85%.
- Utilized CI/CD pipelines and MLOps techniques with Wandb to streamline model development and deployment.
- Collaborated with team to present and won the People's Choice Award at the 2024 Undergraduate Research Expo.

### VinBigData | Vietnam's Leading Multi-Cognitive AI StartUp

Hanoi, Vietnam

*Data Science Intern*

*05/2023 – 08/2023*

- Optimized a LLM question classifier using PyTorch and Transformers, achieving 87% accuracy on 100,000+ questions.
- Expanded the SQuAD 2.0 dataset with 50,000 high-quality question-answer pairs, boosting the accuracy of the Transformers DeBERTA question-answering model by 4-10%.
- Streamlined ML model post-processing by deploying a multilingual semantic analysis model on Torchserve, improving model accuracy checking speed by 200%.

### MBBank

Hanoi, Vietnam

*Data Science Intern*

*06/2022 – 08/2022*

- Engineered a data pipeline to crawl data from 7 real estate websites to create a database of 60,000 entries.
- Developed and fine-tuned a real estate price prediction model with XGBoost, LightGBM, and CatBoost, achieving an error score of 0.17.
- Built a user-friendly web application for house price prediction using Flask, Python, CSS, and HTML.

### FPT Software

Hanoi, Vietnam

*Data Analyst Intern*

*06/2021 – 08/2021*

- Developed an automated data pipeline to extract tasks for a team of 20 people from Trello and Microsoft Task to Microsoft Azure, significantly improving work progress monitoring.
- Integrated the Microsoft Azure database with PowerBI via SQL to create a work performance dashboard displaying overall team productivity and individual contributions, boosting team performance.

## PROJECTS

### DoubletDetection – Biomedical Data Science Hackathon, University of Rochester ([link](#))

*Team Leader, First Place, Undergraduate Division*

*08/2023*

- Led the team to implement and optimize a Machine Learning model in Python to predict doublet cells in RNA data, achieving the highest accuracy score in the Undergraduate Division.