

Hua Yang

☎ (+1) 984-373-9587 | ✉ hyang45@ncsu.edu | in Hua Yang | 🌐 HuaYang.me

Summary

Machine Learning Engineer with strong experience in **large-scale data processing**, **feature engineering**, and **end-to-end ML pipeline development**. Skilled in **cleaning, transforming, and analyzing** complex datasets to fine-tune LLMs, and deploying AI solutions with Python, PyTorch, CUDA, Docker, and cloud platforms.

Skills

Programming Languages: Python, Java, JavaScript, C#, HTML, CSS
ML/AI Frameworks & Tools: PyTorch, Hugging Face, CUDA, Scikit-learn, XGBoost
Software Engineering & DevOps: Git, CI/CD, Docker, AWS, REST API, Flask, Bootstrap
Databases & Operating Systems: MySQL, Oracle, SQLite, Windows, Ubuntu
Networking: TCP/IP, DNS, HTTP, SSH, Socket Programming

Work Experience

Machine Learning Intern

AI Large Model for Intelligent Cognition Center, Pazhou Lab (Guangzhou) May 2022 – May 2023

- Built a sleep disorder detection system, achieving 5/171 nationally and securing a \$10,000 award.
- Improved model accuracy (55% → 59%) across 7 datasets while achieving a 110× runtime speedup.

Projects

Causal Study on Membership Inference Attacks in Code LLMs Sep 2024 – May 2025

- **Tech:** Python, PyTorch, Hugging Face, CUDA
- Evaluated code transformations against membership inference (MI) attacks.
- Fine-tuned CodeGPT, CodeGen, and CodeLlama on transformed datasets with minimal accuracy loss (<1.5%) and significant MI reduction (up to 10.19%).

Causal Analysis of PII Leakage in Code LLMs May 2025 – Sep 2025

- **Tech:** LoRA, OpenAI API, Python
- Revealed heterogeneous privacy risks through training-dynamics and causal analysis.
- Built a real-world PII dataset and fine-tuned representative LLMs.

Movie Recommendation Web App Integrating LLMs Sep 2024 – Dec 2024

- **Tech:** CSS, HTML, CI/CD, SVM, CNN, XGBoost, MySQL
- Led a 4-person team to build a recommendation app, mentoring junior developers.
- Designed and optimized ML models and automated deployment with CI/CD.

Education

NC State University, Raleigh, NC

Ph.D., Computer Science, Advisor: Dr. Bowen Xu

GPA: 3.7/4.0

Aug 2024 – Dec 2027

South China University of Technology, China

M.S., Computer Science

GPA: 4.0/4.0

Sep 2021 – Jun 2024

Central South University, China

B.S., Electronic Engineering

GPA: 4.0/4.0

Sep 2017 – Jun 2021

Publications

- **Yang, H.**, et al. *A Causal Perspective on the Role of Training Dynamics for Interpreting Privacy Risks in Code Models Submitted to FSE 2026* (under review, **top-2** conference in software systems).
- **Yang, H.**, et al. *How Do Semantically Equivalent Code Transformations Impact Membership Inference on LLMs for Code? Submitted to ICSE 2026* (under review, **top-1** conference in software systems).
- **Yang, H.**, Chen, C. P., Chen, B., & Zhang, T. (2024). *Improving the Interpretability through Maximizing Mutual Information for EEG Emotion Recognition*. *IEEE Transactions on Affective Computing*. (Top Journal in AI for Science).
- **Yang, H.**, Chen, C. P., Chen, B., & Zhang, T. (2023). *Faceexplainer: Generating model-faithful explanations for graph neural networks guided by spatial information*. In *2023 IEEE BIBM* (pp. 718-725). IEEE. (Acceptance Rate: 19.5%, Explainable AI).