

# Hui (Amelia) DAI

Email: [hd2584@nyu.edu](mailto:hd2584@nyu.edu) ▪ Website: [ameliadai.github.io](https://ameliadai.github.io)

## EDUCATION

**New York University | Center for Data Science** 05/2025 (Expected)  
**Major:** *Master of Science in Data Science* | **GPA:** 4.0/4.0 New York, NY  
**Core Courses:** Natural Language Understanding, Computer Vision, Machine Learning, Big Data

**The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen)** 09/2019 – 05/2023  
**Major:** *Bachelor of Science in Statistics* | **Major GPA:** 3.928/4.0 (rank: 3/181) Shenzhen, China  
**Core Courses:** Machine Learning (Python, R), Data Structures (Java), Optimization (MATLAB), Regression Analysis, Time Series, Stochastic Processes, Statistical Inference, Probability and Statistics, Calculus, Linear Algebra

## PUBLICATION

- [Hui Dai](#), Ryan Teehan, and Mengye Ren. *Are LLMs Prescient? A Continuous Evaluation using Daily News as the Oracle*. In *NeurIPS Workshop on Adaptive Foundation Models* [Oral], 2024. (Under review at ICLR 2025).
- Renjie Li, Ceyao Zhang, Wentao Xie, ..., [Hui Dai](#), ..., and Zhaoyu Zhang. *Deep reinforcement learning empowers automated inverse design and optimization of photonic crystals for nanoscale laser cavities*. In *Nanophotonics*, 2023.

## RESEARCH EXPERIENCE

**Agentic Learning AI Lab** 03/2024 - Present  
*Supervised by Prof. Mengye Ren* New York, NY  
**Purpose:** To build a continuous evaluation benchmark assessing LLMs' temporal generalization and forecasting abilities  
**Workload:**

- Designed and built an automatic QA generation framework using daily scraped news
- Evaluated and analyzed LLMs' performance on our benchmark, including the RAG experimental setting

**Achievement:** First benchmark on future event prediction that can capture and quantify model performance degradation, calling for continuous pre-training of LLMs to address outdated representations [\[paper\]](#) [\[website\]](#)

**NYU Langone Health** 06/2024 - Present  
*Supervised by Prof. Krzysztof J. Geras* New York, NY  
**Purpose:** To predict breast cancer risk by leveraging multimodal and longitudinal images in survival analysis framework  
**Workload:**

- Extended the Cox proportional hazards model with transformers, integrating various imaging modalities and clinical risk factors to predict 1-5 year breast cancer risk
- Implemented the discrimination and calibration metrics for survival analysis evaluation

**Future work:** Explore advanced time-embedding techniques, address challenges related to high censoring rates, and improve model calibration for more reliable results

**Shenzhen Research Institute of Big Data** 03/2022 – 10/2022  
*Supervised by Prof. Zhaoyu Zhang & Dr. Renjie Li* Shenzhen, China  
**Purpose:** To autonomously learn a certain inverse design of nanophotonic laser cavities using Reinforcement Learning  
**Workload:**

- Debugged and fixed errors in the customized Reinforcement Learning algorithms (PPO and DQN) under a nanophotonic environment with open-source Python libraries *OpenAI gym* and *Ray Rllib*
- Tested the algorithm, tuned hyperparameters, and recorded catalog contents

**Achievement:** Our algorithm raised sampling efficiency over 200 times and achieved 10 times better performance than human experts [\[paper\]](#) [\[code\]](#)

## PROJECT EXPERIENCE

**DENIAHL: In-Context Features Influence LLM Needle-In-A-Haystack (NIAH) Abilities** [\[paper\]](#) [\[code\]](#) Spring 2024  
*Instructor: Sophie Hao, New York University*

**Purpose:** To analyze factors beyond context length affecting LLMs' abilities to recall information from long input context  
**Workload:**

- Created synthetic dataset to ablate NIAH features beyond typical context length including data size, pattern, and type
- Designed and ran experiments to assess changes in model recency biases based on different data input

**Achievement:** Identified that data size impacts performance globally, local attention aids recall with altered patterns, and changing data type from numbers to letters shifts "lost-in-the-middle" (Liu et al., 2023) to "lost-in-the-end" phenomena

## Revisiting the Text-Based Ideal Point Model (TBIP) [\[paper\]](#) [\[code\]](#)

Spring 2024

*Instructor: Rajesh Ranganath, New York University*

**Purpose:** To evaluate TBIP (Vafa et al., 2020), a probabilistic topic model, for estimating ideologies using word-choice differences on shared topics, on its application from biparty to multiparty contexts

### **Workload:**

- Adapted TBIP to analyze over 200k parliamentary speeches (2016–2019) from the UK House of Commons, identifying key latent topics and ideological framing, and exploring modeling ideal points in higher dimension
- Benchmarked TBIP against vote-based ideal points, revealing moderate correlations and challenges in distinguishing ideologies in different parties during the Brexit era

**Achievement:** Demonstrated TBIP’s ability and limitations to infer political preferences in multiparty contexts

## INDUSTRY EXPERIENCE

---

### **Donson Times Technology (Shenzhen) Co., Ltd**

06/2023 – 08/2023

*Machine Learning Engineer Intern*

Shenzhen, China

**Purpose:** To finetune LLMs (ChatGLM and Qwen) for automated marketing copy generation and sentiment identification

### **Workload:**

- Called GPT’s API to identify marketing copy, extract keywords from text data, and design a scoring model for assessing marketing copy quality; proposed a keyword highlighting mechanism for efficient annotation
- Performed prompt tuning, parameter efficient tuning (prefix tuning, LoRA), and explored RLHF techniques
- Built an evaluation pipeline to test fine-tuned models, incorporating single/multi-round evaluations with prompt tuning and few-shot learning

**Achievement:** Our model was put into commercial usage as part of a marketing-focused chatbox

## TEACHING EXPERIENCE

---

### **New York University, Courant Institute of Mathematical Sciences**

*Recitation Leader*

- Mathematics for Economics I (MATH-UA 131) with *Prof. Raoul Normand* Fall 2024
- Mathematics for Economics II (MATH-UA 132) with *Prof. Raoul Normand* Spring 2023
- Linear Algebra (MATH-UA 140) with *Prof. Andrew Sanfratello* Fall 2023

**Responsibilities:** Conduct weekly recitations, grade quizzes via Gradescope, hold office hours, and proctor exams

## SKILLS & OTHERS

---

**Programming:** Python (PyTorch, Scikit-learn, Pandas), R, C++, Java, Stata, MATLAB, SQL, LaTeX, Linux, Slurm

**Awards:** Academic Performance Scholarship, 3 years (2019-2022); Bowen Scholarship, 4 years (2019-2023); Dean's List, 4 years (2019-2024); Second Prize of Guangdong Competition Area in China Undergraduate Mathematical Contest in Modeling (2020)