Mimansa Jaiswal

Research Scientist

mimansa.jaiswal@gmail.com ◆ (734) 747-0283 ◆ mimansajaiswal.github.io ◆ MimansaJ ◆ mimansajaiswal ◆ mimansa@umich.edu

EXPERIENCE

Meta (MSL Org)

Research Scientist   Feb 2025 – Present

- Part of Meta AI's Capabilities team, developed and fine-tuned large language models (LLMs) using advanced post-training techniques including RLHF/RLVR/RLAIF (with and without checklists), and used DPO/GRPO with reward models for improving production models.
- Built end-to-end synthetic data generation pipelines and automated evaluation systems, including rubric-based judges, preference data collection workflows, and quality assessment frameworks for model improvement and hill-climbing optimization.

Norm AI

Research Engineer   Oct 2023 – Oct 2024

- Developed domain-specific retrieval-augmented generation (RAG) systems for legal applications, implementing synthetic data augmentation techniques to expand training datasets, create evaluation datasets and improve model coverage.
- Built agentic LLM simulation frameworks to model human group perception and consensus-building for subjective law interpretation and policy evaluation.

Allen NLP, Allen Institute for AI

Research Intern   Sep 2021 – Dec 2021

- Developed interpretable and decompositional evaluation frameworks for GPT-3 and other language models, creating benchmarks for multi-aspect quality assessment.
- Used the designed evaluation methodologies to measure fine-grained model capabilities and failure modes in natural language understanding tasks.

NLP Team, Facebook AI Research (FAIR)

Research Intern   May 2021 – Aug 2021

- Identified and analyzed systematic failure patterns in Natural Language Inference (NLI) through adversarial testing and error analysis.
- Developed correction strategies and data augmentation techniques to improve model robustness on challenging NLI examples performing experimentation on CTRL model.

Conversation AI, Facebook Research

Research Intern   May 2020 – Aug 2020

- Developed interpretable user satisfaction metrics for conversational AI systems by integrating human knowledge and behavioral signals, and trained weakly supervised hierarchal label models.
- Created generalizable evaluation frameworks combining quantitative metrics with qualitative human feedback for dialogue quality assessment.

EDUCATION

University of Michigan   *Ann Arbor, MI*

Ph.D. in Computer Science and Engineering   Sep 2017 – Aug 2023

- **Honors**: Awarded Barbour Fellowship Amongst Over 1k+ Applicants
- Computational Human Analysis and Integration (CHAI) Lab
- *Advisor*: Prof. Emily Provost

University of Michigan   *Ann Arbor, MI*

M.S. in Computer Science and Engineering   Sep 2017 – May 2019

- **Honors**: GPA: 3.87/4.00
- **Coursework**: Natural Language Processing, Advanced Artificial Intelligence, Bayesian Inference

Institute of Engineering and Technology   *Indore, India*

B.E. in Computer Science Engineering   Jul 2013 – May 2017

- **Honors**: GPA: 3.7/4.00
- **Coursework**: Information Retrieval, Machine Learning, Algorithm Design and Analysis

SELECTED PUBLICATIONS

Privacy Enhanced Multimodal Neural Representations for Emotion Recognition. Mimansa Jaiswal, Emily Mower Provost. *AAAI Conference on Artificial Intelligence (AAAI) 2020*, Feb 2020.

Controlling for Confounders in Multimodal Emotion Classification via Adversarial Learning. Mimansa Jaiswal, Zakaria Aldeneh, Emily Mower Provost. *International Conference on Multimodal Interaction (ICMI) 2019*, Oct 2019.

HONORS AND AWARDS

Barbour Fellowship, *University of Michigan Rackham Graduate School*   Ann Arbor, MI | Sep 2022

- Awarded amongst over 1,000+ applicants

Graduate Student Instructor for Affective Computing Course, *University of Michigan*   Ann Arbor, MI | Jan 2020

Student Representative in CSE Faculty Hiring Committee, *University of Michigan*   Ann Arbor, MI | Sep 2020
CSE Department

Invited Speaker at PyCon 2016, *PyCon Singapore*   Singapore | Jun 2016

Invited Speaker at PyCon 2016, *PyCon India*   India | Oct 2016

National Talent Search Examination Scholarship, *National Council of Educational Research and Training*   India
(NCERT)

- Top 0.01% in India

SKILLS

Research Areas: LLM Post-Training (RLHF/RLVR/RLAIF, DPO/GRPO), Model Evaluation & Benchmarking, Failure Analysis, Synthetic Data Generation, Human-in-the-Loop ML, Retrieval-Augmented Generation (RAG), Natural Language Processing, Multimodal Learning

Technical Skills: Python, PyTorch, Transformers, Machine Learning, Deep Learning, Statistical Analysis, Experimental Design

Professional Activities: Reviewing: ICDMW, ICMI, ACII, CHI, CSCW, *ACL, AAAI, NeurIPS (2019-present), Public Speaking, Technical Writing, Mentoring