Nikhil Mehta

nikhilmehtahockey@gmail.com ❖ (408) 712-1003 ❖ Milpitas, CA ❖ GitHub ❖ Google Scholar ❖ hockeybro12.github.io

EDUCATION

Purdue University   **West Lafayette, IN**
Ph.D. Computer Science (Advisor: Dan Goldwasser, Focus: Natural Language Processing)   *October 2024*

SELECTED WORK EXPERIENCE

Meta SuperIntelligence Labs (MSL)   **Menlo Park, CA**
Senior Research Scientist   *June 2025-Present*
Research Scientist   *May 2024 - June 2024*

- Part of small, core post-training for Llama multimodal (image + text) models, work in Generative AI Research.
    - Post-trained Llama 4, Llama 3.2 Multimodal. Improved performance in accuracy, reasoning, chat
    - Implement post-training, Reward Model, Synthetic Data Generation (RJS), grader, and evaluation infra

Purdue NLP Lab   *West Lafayette, IN*
NLP/AI Student Researcher   *August 2017 - May 2024*

- Large Language Models, Interactive Learning, and Computational Social Science:
    - Developed methods for humans to interact with AI agents, that the agents can learn from. My approach improves the performance of LLMs (RLHF), Graph Neural Networks (GNNs), and other ML models.

Microsoft Research (MSR)   **Redmond, WA**
Research Intern   *May – August 2022*

- Interactive learning for multi-modal tasks, incorporating feedback from humans to improve perf. AACL 2024

Amazon Science   **Boston, MA**
Applied Scientist Intern   *May - August 2020, May – August 2021*

- Predict product images given user queries (NeurIPS InterNLP workshop). NER improvement (AMZN conf.)

Fujitsu Laboratories of America   **Sunnyvale, CA**
Machine Learning Research Intern   *May – August 2019*

- Created a patented approach to predict a relevant joke for a chat-bot to reply with, given a conversation history.

SELECTED PUBLICATIONS

The Llama 3 Herd of Models
Contributor

Using RL to Identify Divisive Perspectives Improves LLMs Abilities to Identify Communities on Social Media
EMNLP 2024 Findings: Nikhil Mehta and Dan Goldwasser

An Interactive Framework for Profiling News Media Sources
NAACL 2024 Nikhil Mehta and Dan Goldwasser

Interactively Learning Social Media Representations Improves News Source Factuality Detection
Findings of AACL 2023: Nikhil Mehta and Dan Goldwasser

Tackling Fake News Detection by Continually Improving Social Context Representations
ACL 2022: Nikhil Mehta, Maria Pacheco, and Dan Goldwasser

Improving Natural Language Interaction with Robots Using Advice
NAACL 2019: Nikhil Mehta and Dan Goldwasser

Improving Grounded Language Understanding by Interacting with Agents Through Help Feedback
Findings of EACL 2024: Nikhil Mehta, Milagro Teruel, Patricio Sanz, Xin Deng, Ahmed Awadallah, Julia Kiseleva

SKILLS + REVIEWING

ML/AI, NLP, Reinforcement Learning, CV, Python, C, Large Scale AI Systems, GRPO, Data, PyTorch, Foundation Models
Conference Reviewing: ACL ARR, ACL, AACL, AAAI, NAACL, EMNLP, COLING, EACL