

# Ishaan Watts

 [Pittsburgh, PA](#)  [+1 \(412\) 925-7734](#)  [iwatts@andrew.cmu.edu](mailto:iwatts@andrew.cmu.edu)  [Google Scholar](#)  [Homepage](#)  [LinkedIn](#)

## EDUCATION

**Carnegie Mellon University (CMU)** – Master of Science

**Major:** Machine Learning

August 2025 - Present

Pittsburgh, PA

**Indian Institute of Technology Delhi (IIT-D)** – Bachelor of Technology

**Major:** Engineering Physics; **Minor:** Computer Science; **CGPA:** 9/10

July 2019 - May 2023

New Delhi, India

## EXPERIENCE

**Google DeepMind**

Pre-Doctoral Researcher with [Dr. Partha Talukdar](#)

Projects: Gemini 2.5-Pro Contributor, Multi-Model Composition, Model Stitching vs. Adapters for Continual Learning

July 2024 - July 2025

Bengaluru, India

**Microsoft Research**

Research Intern with [Dr. Sunayana Sitaram](#)

Projects: Indic LLM Evaluation Leaderboard 🏆, Multicultural Toxicity Detection Dataset 🌐, Multilingual Finetuning and

Benchmarking of LLMs, Education AI-Bot 📖

May 2023 - June 2024

Bengaluru, India

**Torch Investment Management**

Machine Learning Engineer Intern with [Amit Sharma](#)

Project: Stock Price Modelling using Deep Learning

September 2022 - December 2022

Remote / Noida, India

**Udaan**

Data Scientist Intern with [Pranjal Singh](#)

Project: Holistic User-Embeddings using Graph Neural Networks

May 2022 - July 2022

Remote / Bengaluru, India

**Griffith University**

Research Intern with [Prof. Saiful Islam](#)

Project: Malware Detection using Machine Learning

May 2021 - July 2021

Remote / Queensland, Australia

## PUBLICATIONS

C=CONFERENCE, P=PREPRINT

[P1] **Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation**

**Agentic Capabilities** 📄 📖

*Gemini Team, Google DeepMind (including [Ishaan Watts](#))*

[Tech Report]

[C.4] **RTP-LX: Can LLMs Evaluate Toxicity in Multilingual Scenarios?** 📄 📖 🌐

*Adrian de Wynter, [Ishaan Watts](#), et al.*

*AAAI 2025 Artificial Intelligence for Social Impact Track*

[AAAI 2025]

[C.3] **PARIKSHA: A Large-Scale Investigation of Human-LLM Evaluator Agreement on Multilingual and Multi-Cultural**

**Data** 📄 📖 🌐

*[Ishaan Watts](#), Varun Gumma, Aditya Yadavalli, Vivek Seshadri, Swami Manohar, Sunayana Sitaram*

*Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*

[EMNLP 2024]

[C.2] **MAPLE: Multilingual Evaluation of Parameter Efficient Finetuning of LLMs** 📄 📖

*Divyanshu Aggarwal\*, Ashutosh Sathe\*, [Ishaan Watts](#), and Sunayana Sitaram*

*Findings of the Association for Computational Linguistics: ACL 2024*

[ACL Findings 2024]

[C.1] **MEGAVERSE: Benchmarking LLMs across languages, modalities, models and tasks** 📄 📖

*Sanchit Ahuja, Divyanshu Aggarwal, Varun Gumma, [Ishaan Watts](#), et al.*

*Proceedings of the 2024 Conference of the North American Chapter of ACL*

[NAACL 2024]

## HONORS & AWARDS

- **Narotam Sekhsaria Fellowship (\$40,000)** and **K.C. Mahindra Fellowship (\$6,000)** to pursue graduate studies at CMU
- Received **Peer Bonuses** at Google DeepMind for productive collaborations • **6th position** in Amazon ML Challenge 2023
- Bagged **1st position** in Topic Challenge & **Honors** in Executive Challenge at Microsoft Global Hackathon 2024
- **3x IIT-D Semester Merit Awards** for being top 3 in department • **KVPY Fellowship** by Government of India, ranking top 2%
- **Rank 1635** in JEE-Advanced & **99.94 percentile** in JEE-MAINS among 1.4 million candidates to pursue Bachelors at IIT-D

SELECTED PROJECTS

- Model Composition Advances** July 2024 - July 2025  
Advisors: [Dr. Prateek Jain](#), [Dr. Partha Talukdar](#), [Dr. Marc'Aurelio Ranzato](#) Google DeepMind
  - Explored composition of generalist & specialist LLM for **continual domain adaptation & compositional task** (code × i18n)
  - Improved **CALM** architecture & **beat LoRA** under varied budget constraints; presented poster at **Google Research Week '25**
  - Designed a **routing-based composition** with dynamic inference-time expert selection, beating a 2× larger monolithic model
  - Engineered **JAX-based** training, evaluation, and bulk inference pipelines for composed models within the **Gemini codebase**
- Multilingual and Multicultural Evaluation of LLMs** May 2023 - June 2024  
Advisors: [Dr. Sunayana Sitaram](#), [Dr. Kailka Bali](#), [Dr. Vivek Seshadri](#), [Dr. Swami Manohar](#) Microsoft Research
  - Built fair LLM evaluation platform with **People+AI** & **Karya Inc.**, using culturally-nuanced prompt in 10 Indic languages [\[C.3\]](#)
  - Created leaderboard from **120k+ human/LLM evaluation** & analyzed human-LLM agreement and biases across languages
  - Studied **QLoRA finetuning** across 90+ model configs to analyze cross-lingual transfer, forgetting, and PEFT effects [\[C.2\]](#)
  - Benchmarked multilingual capabilities and test-dataset contamination of 20+ LLMs across **23 tasks** and **83 languages** [\[C.1\]](#)
- Culturally-Localized Toxicity Detection Dataset** December 2023 - April 2024  
Advisor: [Dr. Adrian de Wynter](#) Microsoft Redmond
  - Released toxicity benchmark in **38 languages** to support safer LLM deployment, evaluating LLMs across **8 harm categories**
  - Found LLMs struggle with subtle biases (microaggressions) and show performance drops in low-resource languages [\[C.4\]](#)
- Educational AI Web App** June 2023 - November 2023  
Advisors: [Dr. Akshay Nambi](#), [Dr. Tanuja Ganu](#) Microsoft Research
  - Engineered **Shiksha Copilot**, a GenAI web app, featuring a Streamlit frontend and structured telemetry for data analysis
  - Partnered with **Shikshana Foundation** to deploy it in **10+ schools**, enabling teachers to create engaging classroom content
  - 📰 **Media Coverage:** [Satya Nadella Talk](#), [Times of India](#), [The Indian Express](#), [Analytics India Magazine](#), [Microsoft](#)
- Holistic User-Embedding Generation Framework** May 2022 - July 2022  
Mentors: [Pranjal Singh](#), [Mohit Kumar](#) Udaan
  - Generated user-vectors from buyer-seller interaction graphs using multi-relational, multi-entity **hetero-graph autoencoder**.
  - Integrated embeddings into **fraud detection** models, reducing Udaan dispute loss by **2.45%** and secured a full-time offer.

TECHNICAL SKILLS

- Languages:** Python, Java, SQL, Bash, C, C++,  $\LaTeX$ , JAX    • **Tools:** Git, Docker, ROOT, RAG, Autodesk Inventor
- Libraries:** PyTorch, TensorFlow, Numpy, Pandas, Matplotlib, Scikit-Learn, Snsrape, Streamlit, Langchain, Hugging Face
- Advanced Coursework:** Advance Machine Learning, Probability & Statistics, Deep Learning Systems, Algorithms, Calculus, Digital Electronics, Social Computing, Computer Networks, Linear Algebra & Differential Equations, Signals & Systems

VOLUNTEERING

- Mentored six freshmen for their academic and mental wellbeing through **BSW Student Mentorship Program** at IIT Delhi
- Prepared Mental Health Care-Packages comprising poems & articles under **National Service Scheme** during the pandemic
- Worked in the social media promotion department to spread awareness for the **non-profit organization 'No Help Too Big'**

CO-CURRICULAR ACTIVITIES

- Organised and participated in India's **First Model Builder Summit** at the Bangalore Convention Centre with People+AI
- Delivered an invited **Talk** at the **National Research Council Canada** highlighting "Inclusivity in LLM Evaluations" in 2024
- Presented posters at ACL 2024 in Bangkok & EMNLP 2024 in Miami and a regular **Reviewer for ACL-ARR and AAAI**
- Attended the advanced **Winter School on Deep Learning** organised by the Indian Statistical Institute, Kolkata in Spring 2023
- Completed **A1-level in Spanish** as a non-credited course of the Humanities and Social Sciences Department at IIT Delhi

REFERENCES

• Dr. Partha Talukdar ..... Senior Staff Research Scientist, Google DeepMind 🌐

• Dr. Sunayana Sitaram ..... Principal Researcher, Microsoft Research 🌐

• Dr. Adrian de Wynter ..... Principal Applied Scientist, Microsoft Redmond 🌐

• Prof. Abhishek Iyer ..... Assistant Professor, IIT Delhi 🌐