

# 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











July 2019 - May 2023

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

New Delhi, India

## 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 '25]
- [C.3] **PARIKSHA: A Large-Scale Investigation of Human-LLM Evaluator Agreement on Multilingual & Multi-Cultural Data**    [Ishaan Watts](#), Varun Gumma, Aditya Yadavalli, Vivek Seshadri, Swami Manohar, Sunayana Sitaram [EMNLP '24]
- [C.2] **MAPLE: Multilingual Evaluation of Parameter Efficient Finetuning of LLMs**   Divyanshu Aggarwal\*, Ashutosh Sathé\*, [Ishaan Watts](#), and Sunayana Sitaram [ACL Findings '24]
- [C.1] **MEGAVERSE: Benchmarking LLMs across languages, modalities, models and tasks**   Sanchit Ahuja, Divyanshu Aggarwal, Varun Gumma, [Ishaan Watts](#), .... Kalika Bali, and Sunayana Sitaram [NAACL '24]

## EXPERIENCE

**Google DeepMind** | Pre-Doctoral Researcher




July 2024 - July 2025

Projects: Gemini 2.5-Pro Contributor [P1], Multi-Model Composition, Model Stitching vs. Adapters for Continual Learning

- Explored composition of generalist & specialist LLMs for **continual domain adaptation & compositional tasks** (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**.

**Microsoft Research** | Research Intern

May 2023 - June 2024

Projects: Toxicity Detection Dataset  [C.4], Indic LLM Evaluation Leaderboard  [C.3], Multilingual Finetuning [C.2] and Benchmarking of LLMs [C.1], Education Bot  [Satya Nadella](#), [Times of India](#)

- Built culturally-nuanced **evaluation framework** in 10 Indic languages, assessing 30+ LLMs via **120k+** human/LLM judgments.
- Released **38-language toxicity benchmark**, evaluating 10 S/LLMs on 8 harm type, exposing failures in detecting subtle biases.
- Studied **QLoRA finetuning** on 90+ models, analyzing cross-lingual transfer and benchmarked 20 LLMs across **83 languages**.
- Developed & deployed **Shiksha Copilot**, GenAI tool helping teachers create educational content, in **100+ schools** in India.

**Torch Investment Management** | Machine Learning Engineer Intern

September 2022 - December 2022

Project: Stock Price Modeling via Deep Learning

- Optimized **LightGBM** model to predict Top 30 US S&P 500 stocks & modeled Saudi market data; scraped Twitter via sncrape.
- Curated feature linking stock price with tweet sentiment by **topic filtering** with BART and **sentiment analysis** using FinBERT.

**Udaan** | Data Scientist Intern

May 2022 - July 2022

Project: Holistic User-Embeddings via Graph Neural Networks

- 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.

## 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, Sncrape, 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

## HONORS & AWARDS

- **Narotam Sekhsaria Fellowship (\$40,000)** and **K.C. Mahindra Fellowship (\$6,000)** to pursue graduate studies at CMU
- **Peer Bonuses** at Google DeepMind • **1st place** in Microsoft Global Hackathon '24 • **6th place** in Amazon ML Challenge '23
- **3x IIT-D Semester Merit Awards** for being top 3 in department • **KVPY Fellowship** by Government of India, ranking top 2%