Ishaan Watts

✉ wattsishaan18@gmail.com ⬡ Github 🎓 Google Scholar 🌐 Homepage in LinkedIn 🐦 Twitter/X

EDUCATION

Indian Institute of Technology Delhi – Bachelor of Technology   *July 2019 - May 2023*
Major: Engineering Physics; **Minor:** Computer Science; **Department Rank:** 4/49   **CGPA:** 8.996/10

Featured Coursework: Data Structures and Algorithms, Digital Electronics, Machine Learning, Special Topics in Computer Applications (Social Computing), Computer Networks, Analysis and Design of Algorithms, Linear Algebra and Differential Equations, Calculus, Probability and Stochastic Processes, Signals and Systems

EXPERIENCE

Google DeepMind   *July 2024 - Present*
Pre-Doctoral Researcher with Dr. Partha Talukdar   Bengaluru, India
Project: Modularity for Continual Learning

Microsoft Research   *May 2023 - June 2024*
Student Researcher with Dr. Sunayana Sitaram   Bengaluru, India
Projects: Indic LLM Evaluation Leaderboard (PARIKSHA), Multilingual & Multicultural Toxicity Detection Dataset (RTP-LX), Analysis of Multilingual Finetuning of LLMs (MAPLE), Multilingual Benchmarking (MEGAVERSE), Shiksha Copilot

Torch Investment Management   *Sept 2022 - Dec 2022*
Machine Learning Engineer Intern with Amit Sharma   Remote / Noida, India
Project: Stock Price Modelling

Udaan   *May 2022 - July 2022*
Data Scientist Intern with Pranjal Singh   Remote / Bengaluru, India
Project: Holistic User-Embeddings via GNNs

Griffith University   *May 2021 - July 2021*
Research Intern with Prof. Saiful Islam   Remote / Queensland, Australia
Project: Malware Detection using Machine Learning

PUBLICATIONS   C=CONFERENCE, *=EQUAL CONTRIBUTION

[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*, Ashutosh Sathe, Millicent Ochieng, Rishav Hada, Prachi Jain, Maxamed Axmed, Kalika Bali, and Sunayana Sitaram
Proceedings of the 2024 Conference of the North American Chapter of ACL   **[NAACL 2024]**

MAJOR AWARDS & ACHIEVEMENTS

- Bagged **1st position** in Topic Challenge & Honours in Executive Challenge at Microsoft Global Hackathon.   *2023*
- Secured **6th position amongst 5000** teams all over India in the Amazon ML Challenge 2023.   *2023*
- Granted **Merit Award** in 3 semesters for ranking top 3 of 49 students in Physics Dept., IIT Delhi.   *2021-2022*
- Achieved **Rank 1635 in JEE-Advanced** & **99.94 percentile in JEE-MAINS** among 1.4 million candidates.   *2019*
- Recipient of **KVPY Fellowship** (Kishore Vaigyanik Protsahan Yojana) given by DST, Government of India.   *2018*

Featured Academic Projects & Collaborations

- **Modularity for Continual Learning**   *July 2024 - Present*
  Advisors: Dr. Prateek Jain, Dr. Partha Talukdar   Google DeepMind
  - Experimenting different **model composition** architectures in LLMs to continually adapt models to new domain.
  - Analyzed continual learning methods across training flops, added parameters, token budget and inference latency.
  - Curated **composition tasks** like **code x i18n** & presented findings as poster at Google Research Week **[GRW'25]**.

- **Indic LLM Evaluation Leaderboard**   *Jan 2024 - June 2024*
  Advisors: Dr. Sunayana Sitaram, Dr. Vivek Seshadri, Dr. Swami Manohar   Microsoft Research India
  - Collaborated with People+AI and Karya Inc. to build a fair and transparent evaluation platform for LLMs. 🏆
  - Curated **culturally-nuanced** prompt through native-speakers in **10 Indian languages** and evaluated 30+ LLMs.
  - Built leaderboards from **90k human** & **30k LLM-based** evaluations via pairwise comparisons & direct assessment.
  - Studied Human-LLM agreement and potential biases in evaluations across languages and models. **[EMNLP'24]**

- **Multicultural and Multilingual Toxicity Dataset Curation**   *Dec 2023 - April 2024*
  Advisor: Dr. Adrian de Wynter   Microsoft Redmond
  - Released culturally-nuanced toxic multilingual benchmark in **38 languages** to facilitate safer deployment of LLMs.
  - Evaluated **10 S/LLMs** for toxicity detection across **8 harm categories** and different language groups. **[AAAI'25]**
  - Observed LLMs fail to detect subtle biases (microagressions) & performance decreases with language's availability.

- **Multilingual Finetuning and Benchmarking**   *Oct 2023 - Feb 2024*
  Advisors: Dr. Sunayana Sitaram, Dr. Kailka Bali   Microsoft Research India
  - Studied **QLoRA** finetuning technique across different rank-quantisation configurations, base models & datasets.
  - Finetuned **90+ models** and analysed cross-lingual transfer, catastrophic forgetting and effects of **PEFT**. **[ACL'24]**
  - Benchmarked multilingual capabilities & contamination of 20 LLM across 23 dataset & 83 languages. **[NAACL'24]**

- **Shiksha Copilot Web App**   *June 2023 - Nov 2023*
  Advisors: Dr. Akshay Nambi, Dr. Tanuja Ganu   Microsoft Research India
  - Engineered **Shiksha Copilot** web app to aid teachers in creating engaging content for students using Gen-AI.
  - Built an intricate and user-friendly frontend via streamlit, structured telemetry and analysed collected data.
  - Collaborated with Shikshana Foundation and deployed Copilot in **10+ schools** across Karnataka to teachers.
  - 📰 **Media Coverage:** *Satya Nadella talk, Times of India, The Indian Express, Analytics India Magazine, Microsoft*

- **Holistic User-Embeddings via GNNs**   *May 2022 - July 2022*
  Mentors: Pranjal Singh, Mohit Kumar   Udaan
  - Developed framework to generate user representation from buyer-seller interaction graph for better segmentation.
  - Built multi-relational & multi-entity graph and modeled **Hetero-Graph AutoEncoder** with novel a loss function.
  - Decreased Udaan dispute loss by **2.45%** by incorporating embeddings in existing fraud model – **full-time offer**.

Volunteering & Co-Curricular Activities

- **Co-Curricular Activities**
  - Organised & participated in India's **First Model Builder Summit** at Bangalore Convention Centre with People+AI.
  - Gave a **Talk** promoting "Inclusivity in Evaluations" at National Research Council Canada (NRCC) in Sept'24.
  - Presented papers at ACL'24 & EMNLP'24 and regular **Reviewer for ARR** (June'24, August'24 & October'24 cycles).
  - Attended **Winter School on Deep Learning** organised by Indian Statistical Institute (ISI), Kolkata in Spring 2023.
  - Completed **A1 level in Spanish** as non-credited course of Humanities & Social Sciences Department, IIT Delhi.

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