

HARMAN SINGH

Pre-doctoral Researcher, Google DeepMind

+919870266967

HarmanDotpy

harmandotpy.github.io

harmansingh.iitd@gmail.com

ACADEMICS

- **Indian Institute of Technology Delhi (IIT-D)**
B.Tech in Electrical Engineering; GPA: 9.349/10.0

July 2018 - June 2022

RESEARCH/WORK EXPERIENCE

- **Google DeepMind**, Pre-doctoral Researcher Aug 2023 - present
Advisor: Dr. Partha Talukdar, Dr. Trevor Cohn, Dr. Sriram Ganapathy
 - **Gemini Core team member:** Developed **data-curation** techniques for improving performance on **low-resource languages**. Building improved evaluations for measuring progress on low-resource language and speech understanding.
 - **Long-context multimodal LLMs:** Benchmarking and improving **in-context learning** capabilities of long-context, multilingual and multimodal (having speech and vision capabilities) models.
 - **IndicGenBench** [1]: Created a high-quality and large-scale benchmark for evaluating LLMs on 29 Indic Languages.
 - **Speech Tokenizer Assessment Benchmark (STAB)** [6]: Developed a cost effective methodology for evaluation of speech-tokenizers without requiring any pre-training.
- **FAIR, Meta AI**, AI Resident July 2022 - July 2023
Advisor: Dr. Pengchuan Zhang, Dr. Hugo Chen, Dr. Wenhan Xiong and Dr. Qifan Wang
 - Improving the compositionality of contrastively pre-trained vision-language models.
 - **MosaiCLIP** [2]: Created a graph decomposition and augmentation framework with coarse-to-fine contrastive learning for improved relation understanding and systematic generalization, while matching CLIP's performance on general tasks.
- **INK Lab USC**, Undergraduate Researcher June 2021 - Jan 2022
Advisor: Prof. Xiang Ren
 - Developed a **causal and modular**, logical reasoning architecture (**FaiRR**) [7] for **deductive reasoning** over natural language rulebases. This method is more interpretable and robust to linguistic variations, compared to previous methods
- **IBM Research AI**, Research Intern July 2021 - June 2022
Advisor: Prof. Parag Singla, DAIR Lab, IIT Delhi and Dr. Dinesh Garg, IBM Research India
 - Worked on creating a Neuro-Symbolic and **Object-Centric models** for multimodal (vision and language) reasoning
 - Developed a weakly supervised text-based image-editing model [5], which can be trained without ground truth output image supervision. Our model offers greater interpretability, and OOD generalization capabilities
- **LimeChat (YC'21)**, NLP Engineering Intern May 2021 - July 2021
 - Built Level 3 AI chatbots (pre-ChatGPT) for D2C brands.
- **Mangul Lab USC**, Undergraduate Bioinformatics Researcher May 2020 - March 2021
Advisor: Prof. Sergei Mangul
 - Designed new algorithms for Phylogenetic analysis of SARS-CoV-2 genomic data, being more accurate and scalable
 - New network representation algorithm paper published in **ISBRA 2021** [9]. Presented at **ABACBS 2020** and **CAME 2020** workshop (ACM-BCB 2020). Contributed to a commentary article [8] published in **Nature Methods**

PUBLICATIONS

Conference Publications

1. **IndicGenBench: A Multilingual Benchmark to Evaluate Generation Capabilities of LLMs on Indic Languages**
Harman Singh, Nitish Gupta, Shikhar Bharadwaj, Dinesh Tewari, Partha Talukdar
ACL 2024 (Main) [[Paper](#)][[Github](#)][[HuggingFace](#)]
2. **Coarse-to-Fine Contrastive Learning in Image-Text-Graph Space for Improved Vision-Language Compositionality**
Harman Singh, Pengchuan Zhang, Qifan Wang, Mengjiao Wang, Wenhan Xiong, Jingfei Du, Yu Chen
EMNLP 2023 (Main) [[Paper](#)]
Oral at **CLVL workshop, ICCV 2023**; and Oral at **SpLU-RoboNLP workshop, EMNLP 2023**
3. **Cross-Lingual Multi-Hop Knowledge Editing**
Harman Singh*, Aditi Khandelwal*, Hengrui Gu, Tianlong Chen, Kaixiong Zhou
EMNLP 2024 (Findings) [[Paper](#)]

4. **Generate-Plan-Edit : A Framework for Improving Image-Text Alignment for Compositional Text-to-Image Generation**
Ashish Goswami, Satyam K. Modi, Santhosh R. Deshineni, Harman Singh, Prathosh A P, Parag Singla
Under Review at CVPR 2025
5. **Image Manipulation via Multi-Hop Instructions - A New Dataset and Weakly-Supervised Neuro-Symbolic Approach**
Harman Singh, Poorva Garg, Mohit Gupta, Kevin Shah, Arnab Kumar Mondal, Dinesh Khandelwal, Parag Singla, Dinesh Garg
EMNLP 2023 (Main) [\[Paper\]](#)
Neuro Causal and Symbolic AI (nCSI) workshop at NeurIPS 2022
6. **STAB: Speech Tokenizer Assessment Benchmark**
Harman Singh*, Shikhar Vashishth*, Shikhar Bharadwaj*, Sriram Ganapathy, Chulayuth Asawaroengchai, Andrew Rosenberg, Kartik Audhkhasi, Ankur Bapna, Bhuvana Ramabhadran
Under Review at ICASSP 2025 [\[Paper\]](#)
7. **FaiRR: Faithful and Robust Deductive Reasoning over Natural Language**
Soumya Sanyal, Harman Singh, Xiang Ren
ACL 2022 (Main) [\[Paper\]](#)[\[Code\]](#)

Journal Publications

8. **Unlocking capacities of viral genomics for the COVID-19 pandemic response**
Sergey Knyazev, Karishma Chhugani, Harman Singh*, Varuni Sarwal*, Ram Ayyala* et al.
Nature Methods [\[Paper\]](#)
9. **A Novel Network Representation of SARS-CoV-2 Sequencing Data**
Sergey Knyazev, Daniel Novikov, Mark Grinshpon, Harman Singh, Ram Ayyala et al.
International Symposium on Bioinformatics Research and Applications 2021 [\[Paper\]](#)[\[Code\]](#)

*equal contribution

KEY PROJECTS

- **IndicGenBench: Evaluation of LLMs on 29 Indic Languages** 📖 🌐 [huggingface](#) Aug 2023 - May 2024
Dr. Partha Talukdar, Dr. Nitish Gupta, Google Deepmind
 - Largest benchmark for Indic evaluation of LLMs in 29 of Indic languages, 13 scripts and 4 language families.
 - 5 user-facing tasks including cross-lingual summarization, multilingual and cross-lingual QnA and Machine Translation.
 - Studied transfer from high-resource languages, deficiencies in tokenization, improvements with in-context learning and that in-context learning is better than FT on thousands of examples.
- **Improving Compositionality of Vision Language Models** 📖 Aug 2022 - Aug 2023
Dr. Pengchuan Zhang, Dr. Hugo Chen, Meta AI
 - Solving critical deficiencies of vision and language models like CLIP, in terms of compositionality
 - Exploring hard negative image and text mining, and data augmentations for improving compositional generalization
 - Experiments include distributed pre-training and finetuning of large scale vision-language models on large image-text paired datasets (20-100 Million)
- **STAB: Speech Tokenizer Assessment Benchmark** Aug 2023 - March 2024
Dr. Partha Talukdar, Ankur Bapna, Dr. Bhuvana Ramabhadran, Google Deepmind
 - Developed STAB, the first low-cost methodology and benchmark for evaluating quality of speech-tokens, without having to pre-train and fine-tune foundation speech or speech-text models.
 - STAB measures various tokenizer properties such as compressibility, robustness and speaker agnosticity etc. which strongly correlate with downstream task performance.
- **Neuro-Symbolic Image Manipulation** 📖 June 2021 - May 2022
Prof. Parag Singla, IITD and Dr. Dinesh Garg, IBM Research
 - Designed NeuroSIM, a **weakly supervised**, modular, neuro-symbolic architecture for text guided image manipulation.
 - The created model is data efficient, interpretable by design, and can generalize to complex text instructions and scenes
- **Faithful and Robust Deductive Reasoning over Natural Language** 🌐 📖 June 2021 - Jan 2022
Prof. Xiang Ren, Ink Lab, USC
 - Designed a 3 step modular architecture for interpretable and robust deductive reasoning over natural language theories containing rules + facts. Generated data for evaluating our model's robustness w.r.t ProofWriter (Tafjord et. al.)
 - Modeled the process of **rule selection**, **fact selection**, **conclusion generation** separately using 3 transformer models (RoBERTa and T5 models). This architecture's proof generation process is causal by design, and it's reasoning steps are more interpretable compared to baselines

- Improved **average robustness** to linguistic perturbations by **2.2%**, and improved **consistency of predictions** by **3%**
- **Solving Visual Combinatorial Problems using Deep Learning** 📄 Jan - May 2021
Prof. Parag Singla, IIT Delhi
 - Implemented a **Visual Sudoku Solver (VSS)** to solve sudoku boards made of handwritten digits of **Arabic MNIST**
 - Used semi-supervised clustering methods like **Unsupervised Data Augmentation (UDA)**, and **InfoGAN** to classify Arabic MNIST dataset using **just 1 labeled sample** per class. Achieved **90%+** clustering accuracy using UDA
 - Implemented a **Recurrent Relational Network (RRN)** for solving a symbolic sudoku. Trained the digits classifier and RRN, end to end and obtained a VSS having **95%+** accuracy of solving an input visual sudoku board
- **Semi-Supervised Conditional GANs** 📄 Jan - March 2021
Self Project
 - Developed and trained the **original cGAN** with limited (~10) labelled + unlabelled samples of MNIST, CIFAR10 etc
 - Used **KMeans** to get noisy clusters and trained the cGAN with these labels. Obtained **FID score = 49.9** on MNIST
 - Implemented **Unsupervised Data Augmentation** for clustering and a linear **label-transforming layer** (to learn from noisy labels) as an inductive bias which improved the **FID score to 45.4** using the original cGAN paper architecture

ACHIEVEMENTS

- **Oral acceptance** at Closing the loop between vision and language workshop at ICCV 2023 [[website link](#)]
- **Oral acceptance** at SpLU-RoboNLP workshop at EMNLP 2023 [[website link](#)]
- **Outstanding Reviewer Award**, MLRC 2022 (Received GCP credits worth \$5k USD)[[website link](#)]
- One of 4 from India and 27 from the world to be selected for the **AI Residency** position at **Meta AI**
- Selected for **Pre-doctoral/Research Fellow** position at **AllenAI, Seattle** and **Microsoft Research, India**, 2022
- Selected for CIFAR Deep Learning Reinforcement Learning Summer School
- **IIT-Delhi Semester Merit Award** for being amongst **top 7%** students in 4 out of 8 semesters at IIT Delhi
- Secured **All India Rank 170** in IIT JEE Mains 2018 and **751** in IIT JEE Advanced amongst 1.3 million candidates
- **KVPY Scholarship** (All India Rank **160**) in 2016-17 by Department of Science and Technology Govt. of India.
- Recipient of **Professor S.K Saha award** for the **best robotics team** in IIT Delhi.
- Placed in **top 0.1%** students in India by **securing 100%** in Mathematics and Computer Science in class 12

ACADEMIC SERVICE AND TEACHING

- **Reviewing:**
 ICLR 2024, COLM 2024, ACL-ARR 2024 (Feb, April, June), NeurIPS 2023, EMNLP 2023, MLRC 2022 (**Outstanding Reviewer Award**)
- **TA for Machine Intelligence and Learning** [✓] Aug - Dec 2021
Instructors: Prof Sumeet Agarwal and Prof Jayadeva
 Responsible for conducting problem solving and programming (python/pytorch) tutorials, grading assignments, taking vivas for projects and course assignments, creating assignment questions.
- **TA for Introduction to Electrical Engineering** Nov 2021 - March 2022
Instructors: Prof Anuj Dhawan
 Responsible for conducting problem solving tutorials, creating exam questions and grading exam copies
- **Demo Leader, NeurIPS Education Outreach Program** 2022
 Gave a demo on compositional generalization in large ML models, to 240+ high school students

SKILLS

- **Languages:** Python, Java, C, C++, MATLAB, Verilog, LaTeX, Bash
- **ML/DL Libraries:** PyTorch(Adv), Tensorflow(Intermediate), HuggingFace(transformers), Scikit-Learn, NLTK, Spacy
- **Tools:** Git, Autodesk Inventor, Quartus, Android Studio, Arduino, Rasberry Pi

RELEVANT COURSEWORK

Machine Learning, Deep Learning(Special Topics in Machine Learning), Meta Learning(Special Module in Machine Learning), Markov Decision Process and Reinforcement Learning, Probability and Stochastic Processes, Linguistics(Language Science), Natural Language Processing, NLP Seminar (Special Module in AI), Data Structures and Algorithms, Analysis and Design of Algorithms, Linear Algebra and Differential Equations, Calculus, Signals and Systems

Online: Stanford CS229-Machine Learning, CS231-Convolution Neural Networks for Visual Recognition, Deep Learning and Neural Networks