Harman Singh

AI Resident, Meta
1 Hacker Way, Menlo Park CA 94025

📞 +12066434953
 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 Research,** Pre-doctoral Researcher   Aug 2023 - present
  Advisor: Dr. Partha Talukdar
  - Research on multimodal, multilingual models with a focus on low resource languages.

- **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.[1]

- **INK Lab USC,** Undergraduate Researcher   June 2021 - Jan 2022
  Advisor: Prof. Xiang Ren
  - Developed a **causal and modular**, logical reasoning architecture (**FaiRR**) [3] for **deductive reasoning** over natural language rulebases. This method is more interpretable and robust to linguistic variations, compared to previous methods
  - Compared FaiRR's logical reasoning capabilities with baseline methods based on language models like BERT, Roberta, T5, through multiple experiments, and empirically proved it to be more robust to linguistic variations

- **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 [2], 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 - July 2021
  - Built and improved Level 3 AI chatbots for D2C brands using various technologies like PostgreSQL, Redis, Rasa etc
  - Implemented **auto-labelling** of conversations including **sentiment analysis, user intent classification etc**
  - Finetuned language models like **BERT, Rasa's DIET classifier** for obtaining text embeddings and assigning labels

- **Mangul Lab USC,** Undergraduate Bioinformatics Researcher   May 2020 - March 2021
  Advisor: Prof. Serghei Mangul
  - Developed a novel way for Phylogenetic analysis (of spread) of Covid-19 using worldwide Covid-19 genomic data
  - Designed new algorithms for creating graphs and trees of SARS-CoV-2 genomic data, being more accurate and scalable
  - Our network representation algorithm paper has been published in **ISBRA 2021** [5]. Presented at **ABACBS 2020** and **CAME 2020** workshop (ACM-BCB 2020). Contributed to a commentary article [4] published in **Nature Methods**

Publications

Conference Publications

1. **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 (long paper, main conference) [Paper]
   Oral acceptance to CLVL Workshop at ICCV 2023

2. **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 (long paper, main conference) [Paper]
   Neuro Causal and Symbolic AI (nCSI) workshop at NeurIPS 2022

3. **FaiRR: Faithful and Robust Deductive Reasoning over Natural Language.**
   Soumya Sanyal, Harman Singh, Xiang Ren
   ACL-2022 (long paper, main conference) [Paper|Code]

Journal Publications

4. **Unlocking capacities of viral genomics for the COVID-19 pandemic response.**
   Sergey Knyazev, Karishma Chhugani, <u>Harman Singh</u>*, Varuni Sarwal*, Ram Ayyala* et al.
   Nature Methods [Paper]

5. **A Novel Network Representation of SARS-CoV-2 Sequencing Data.**
   Sergey Knyazev, Daniel Novikov, Mark Grinshpon, <u>Harman Singh</u>, Ram Ayyala et al.
   International Symposium on Bioinformatics Research and Applications 2021 [Paper|Code]

*equal contribution

KEY PROJECTS

Deep Learning, NLP, Vision and Language

- **Compositionality Aware Pre-training for Vision Language Models** 📄   Aug 2022 - Present
  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)

- **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

- **An Analysis of Normalizations in Deep Learning for Vision Applications** ⚙ 📄   Jan - Feb 2021
  Prof. Parag Singla, IIT Delhi
    - Implemented **Batch Norm, Instance Norm, Batch-Instance Norm, Layer Norm and Group Norm** from scratch
    - Implemented a generic **ResNet model**. Trained **ResNet-14** on **CIFAR10**, with and without the above normalizations
    - **Model with Batch-Instance Norm** gave best results for classification with **81.1% accuracy, 81.1% MacroF1 score**

- **MultiLingual Question Answering** ⚙   Nov, 2021
  Prof. Mausam, NLP Course Assignment, IIT Delhi
    - Trained a model to learn from **Hindi+Tamil QnA** data, for predicting the answer span in a given context.
    - Finetuned RoBERTa-large on Squad v2 followed by on **chaii-1+MLQA dataset**. Got **70.1%** test accuracy (class-top5)

- **Inducing Constraints in Named Entity Recognition Systems** ⚙ 📄   Feb - March 2021
  Prof Parag Singla, Deep Learning course project, IIT Delhi
    - Developed a Named Entity Recogniton (NER) system for the **GMB** dataset using a **BiLSTM** model with **Layer Norm**
    - Experimented with **random/glove word embeddings, char level embeddings**. Wrote a **CRF module** from scratch and implemented the **Viterbi algorithm**. Got **3.1%** increase in MacroF1 score compared to simple BiLSTM model

Robotics Projects

- **Four Legged Walking Robot with Vision (Minitaur)** 🎧 📄
  Prof Sunil Jha, Prof S K Saha, IIT Delhi   *Dec 2018 - June 2019*
  - Co-Developed a 4-legged robot using 5bar mechanism for its legs.**Achieved turning, slope and step climbing**
  - Developed toe trajectories and gaiting systems for walking including diagonal gate (trotting) and wave gate
  - Integrated Vision using **OpenCV on Rasberry PiCam** for navigating and avoiding obstacles

- **Mecanum Four Wheeled Drive**
  Prof Sunil Jha, Prof S K Saha, IIT Delhi   *Dec 2018 - June 2019*
  - Developed a **4-wheeled Mecanum drive** with **distance sensors, Encoders, IMU (orientation sensors)**
  - Capable of picking and throwing blocks of dimensions approx. 30cm .30cm .30cm up to a distance of 2.2 meters
  - Capable of passing around obstacles, passing a baton, and throwing blocks , as a part of **ROBOCON 2019**

ACHIEVEMENTS

- **Oral acceptance** at Closing the loop between vision and language workshop at ICCV 2023 [website link]

- **Outstanding Reviewer Award**, MLRC 2022 [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:**
  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