Ayush Maheshwari

Senior Solutions Architect, NVIDIA

<firstname>.hakmn@gmail.com • ayushbits.github.io • ayushx
Ayush • ayushbits

Experience

NVIDIA — **Gurugram**
Senior Solutions Architect, NVAITC — *Oct 2024 – Present*

Vizzhy Inc — **Bengaluru**
Lead Research Scientist, Language Model Team — *Sep 2023 – Sep 2024*

1. Led a team of 5 people to build large language models from scratch.
2. Developing data collection & processing pipelines for training and evaluation.
3. Training of tokenizer and designing model training architecture.
4. Training the model on large accelerator cluster.
5. Instruction tuning and preference training of the pre-trained models.

Adobe Research — **Bengaluru**
Research Intern, Document Cloud Team — *May – Aug 2021*

NCETIS, IIT Bombay — **Mumbai**
Project Engineer — *May 2016 – Dec 2018*

Water & Climate Lab, IIT Gandhinagar — **Gandhinagar**
Junior Research Fellow — *May 2015 – Apr 2016*

Tata Consultancy Services — **Mumbai**
Systems Engineer — *Oct 2011 – June 2013*

Education

Indian Institute of Technology Bombay — **Mumbai, India**
Ph.D. in Computer Science and Engineering — *Jan 2019 – Aug 2023*
Defence: July 2024; '**Impactful Research Award 2023**' by IIT Bombay

Thesis: Knowledge Integration in NLP Models using Constraint Ingestion and Generation
Advisor: Prof. Ganesh Ramakrishnan

Tata Institute of Social Sciences — **Mumbai, India**
M.Sc. — *2013 – 2015*

University College of Engineering — **Rajasthan Technical University Kota, India**
B.Tech. (CSE) — *2007 – 2011*

Featured Publications

1. LexGen: Domain-aware Multilingual Lexicon Generation, ACL (Main), 2025
   Maheshwari, A., Singh A., Karthika NJ, Bhat K., Jyothi P., Ramakrishnan G.
2. DictDis: Dictionary Constrained Disambiguation for Improved NMT, Findings of EMNLP, 2024

Maheshwari, A., Jyothi P., Ramakrishnan G.

3. Adaptive mixing of auxiliary losses in supervised learning, AAAI, 2023 (Oral Presentation)
   Sivasubramanian D., **Maheshwari, A.**, Prathosh AP, Shenoy P., Ramakrishnan G.
4. Learning to Robustly Aggregate Labeling Functions for Semi-supervised Data Programming, Findings of ACL, 2022
   Maheshwari, A., Killamsetty K., Ramakrishnan G., Iyer R., Danilevsky M., Popa L.
5. Data Programming using Semi-Supervision and Subset Selection, Findings of ACL, 2021
   Maheshwari, A., Chatterjee O., Killamsetty K., Iyer R., Ramakrishnan G.
6. Joint Learning of Hyperbolic Label Embeddings for Hierarchical Multi-label Classification, EACL 2021
   Maheshwari A., Chatterjee S., Ramakrishnan, G. Jagaralpudi S.
7. UDAAN-Machine Learning based Post-Editing tool for Document Translation, CODS-COMAD Demonstrations, 2023 (Best Demo Paper Award)
   Maheshwari, A., R. Ajay., Subramanian V., Ramakrishnan G.

Open-Source contributions

- **SPEAR** : Data-Programming python library to create training data. GitHub
- **UDAAN** : Post-editing workbench to quickly create publishable translations. Website

Research Interests

- Develop large language models for mid and low-resource languages.
- Develop machine learning based methods that generalize to new domains with limited resources.
- Lexically constrained neural machine translation, data programming, semi-supervised learning

Awards, Services & Talks

- My PhD thesis work was awarded **Impactful Research Award 2023** by IIT Bombay
- Best Demo Paper Award at CODS-COMAD 2023
- Awarded prestigious Ekal Fellowship 2020
- Reviewer: ARR 22–24, EACL, EMNLP, ACL (21–23)
- Fully covered grant for attending EMNLP 2022, Abu Dhabi by EMNLP DEI Committee
- Co-Founder at non-profit Temples of India - A web and app platform listing $>$ 700K temples across Indian subcontinent.
- Key Member of Udaan Project - A project to translate education textbooks in Indian languages.

Computer Skills

Programming Languages: Python, C, C++, Javascript/jQuery, Bash, SQL

Libraries: PyTorch, DeepSpeed, HuggingFace, NumPy, scikit-learn

Talks

- Delivered Masterclass on **Building data-efficient large language models**
  NLTF NASSCOM, 2024
- Invited Tutorial on **Data Efficient Machine Learning for end-to-end Educational Content Creation**
  International Conference on Educational Data Mining, 2023
- Talk on **SPEAR: Data Programming using Semi-Supervision** [Talk]
  DECILE Workshop, 2022